



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 162983

TO: Phuong Bui
Location: REM-2A15&2C18
Art Unit: 1638
Monday, August 29, 2005

Case Serial Number: 10/677179

From: Edward Hart
Location: Biotech-Chem Library
REM-1A55
Phone: 571-272-2512

edward.hart@uspto.gov

Search Notes

Examiner Bui,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

BEST AVAILABLE COPY

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 21:09:59 ; Search time 6876 Seconds
(without alignments)

11458.432 Million cell updates/sec

Title: us-10-677-179-7

Perfect score: 1626

Sequence: 1 gcacgagctcagcttcgtct.....caaaaaaaaaaaaaaaaaa 1626

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1626	100.0	1626	AR432943	AR432943 Sequence
2	877.2	53.9	1607	AK099850	AK099850 Oryza sat
3	870.4	53.5	1635	AR432940	AR432940 Sequence
4	621	38.2	966	AR432945	AR432945 Sequence
5	580.6	35.7	1701	LECHOSYNA	Z21791 L.esculentu
6	574.6	35.3	1802	LECHOSYNA	Z21791 L.esculentu
7	566.4	34.8	1754	CSCHOSYN	X60544 C.embervir
8	535.4	32.9	1311	CO805346	CO805346 Sequence
9	535.4	32.9	1311	BT015785	BT015785 Arabidops
10	535.4	32.9	1603	AY057519	AY057519 Arabidops
11	500.6	30.8	1015	AR432942	AR432942 Sequence
12	500.6	30.8	1057	BT009607	BT009607 Triticum
13	351.4	21.6	659	AB122084	AB122084 Oryza sat
14	351.4	21.6	339650	AP003583	AP003583 Nostoc sp
15	349	21.5	110000	AP006487_1	Continuation (2 of
16	346.2	21.3	1884	SSRPAROC	D67516 Synchocyst
17	346.2	21.3	135551	1 D90906	D90906 Synchocyst
18	340.4	20.9	299350	AP005370	AP005370 Thermosyn
19	316.4	19.5	349742	1 BX572090	BX572090 Prochloro

20	313.2	19.3	599	8	AP457954	AP457954 Zea mays
21	281.2	17.3	306803	1	AE017161	AE017161 Prochloro
22	280.4	17.2	348764	1	BX569689	BX569689 Synchococ
23	279.8	17.2	110000	1	CR522870_34	Continuation (35 o
24	278.6	17.1	110000	8	CR382125_08	Continuation (9 of
25	274	16.9	1053	6	AR236632	AR236632 Sequence
26	274	16.9	1053	6	AR236638	AR236638 Sequence
27	274	16.9	1053	6	AX464573	AX464573 Sequence
28	274	16.9	1053	6	AX464579	AX464579 Sequence
29	262.8	16.2	1876	8	SCYGL148W	Z72670 S.cerevisia
30	262.8	16.2	3036	8	SCAROA	X60190 S.cerevisia
31	262.8	16.2	22803	8	SCV11228	X99960 S.cerevisia
32	261.6	16.1	110000	8	CR382137_03	Continuation (4 of
33	255.4	15.7	10157	1	AE011207	AE011207 Leptospi
34	255.4	15.7	300693	1	AE017287	AE017287 Leptospi
35	254.8	15.7	110000	8	CR380951_4	Continuation (5 of
36	253.2	15.6	348071	1	BX572100	BX572100 Prochloro
37	247.8	15.2	110000	1	AP006841_43	Continuation (44 o
38	239.4	14.7	110000	8	AE016817_01	Continuation (2 of
39	235.2	14.5	301731	1	AE016934	AE016934 Bacteroid
40	231	14.2	110000	8	CR382130_22	Continuation (23 o
41	216.8	13.3	252250	1	AP005339	AP005339 Vibrio vu
42	216.6	13.3	110000	1	BX908798_10	Continuation (11 o
43	213.6	13.1	300045	1	AE016803	AE016803 Vibrio vu
44	203.8	12.5	1080	6	AR548968	AR548968 Sequence
45	199.2	12.3	299180	1	AP005080	AP005080 Vibrio pa

ALIGNMENTS

RESULT 1
AR432943
LOCUS AR432943 Sequence 7 from patent US 6653531. 1626 bp DNA linear PAT 18-DEC-2003
DEFINITION AR432943
ACCESSION AR432943
VERSION AR432943.1 GI:40195584
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1626)
AUTHORS Cahoon,R.E. and Falco,S.C.
TITLE Chorismate synthase from plants
JOURNAL Patent: US 6653531-A 7 25-NOV-2003;
FEATURES
source location/Qualifiers
1..1626
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 1626; DB 6; Length 1626;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCACGAGCTCAGCTTCGCTCTCTCGCGCGCGCGCACTATCATCTTCATTAG 60
DB 1 GCACGAGCTCAGCTTCGCTCTCTCGCGCGCGCGCACTATCATCTTCATTAG 60
QY 61 CTCATCCAACTATTTCGATGACGACCGTGCCCAAGCCACAGCAGGTGGCGCACTCAGG 120
DB 61 CTCATCCAACTATTTCGATGACGACCGTGCCCAAGCCACAGCAGGTGGCGCACTCAGG 120
QY 121 GCACGGCTCGCACCCCGCGCGATCGGCGCTTGTGTGAGTTTGGCCCAAGCTCTCTCTCC 180
DB 121 GCACGGCTCGCACCCCGCGCGATCGGCGCTTGTGTGAGTTTGGCCCAAGCTCTCTCTCC 180
QY 181 CTCGCTTCGCGCGTGACCGCTGCGCACTGCTCGCTTAGAGTTGAGGATCTGGAAC 240
DB 181 CTCGCTTCGCGCGTGACCGCTGCGCACTGCTCGCTTAGAGTTGAGGATCTGGAAC 240
QY 241 ACGTTTGAAACTACTTTTCAGTTTGCAACCTATGTTGAATCTCATGGGGGTGGTTCGT 300

Db 241 ACGTTTGGAACTACTTTTCAGGTTGCAACCTATGTTGAATCTCATGGGGGTGGTGTGGT 300
Qy 301 TGTGTTATCAGTGGTGTCCACCTAGAAATTCACCTCACTGAGGCGAGACCTCAAGTTGAA 360
Db 301 TGTGTTATCAGTGGTGTCCACCTAGAAATTCACCTCACTGAGGCGAGACCTCAAGTTGAA 360
Qy 361 CTCGATCGAAGACGCGCGGACAGACAGAGATAACCTCCACAAGGAAGAGACGTGATACA 420
Db 361 CTCGATCGAAGACGCGCGGACAGACAGAGATAACCTCCACAAGGAAGAGACGTGATACA 420
Qy 421 TGCAAAATTCCTGTCAGGGACACATGAAGGGGTGACTACTGGAACGCCAAATTCCTGTTAT 480
Db 421 TGCAAAATTCCTGTCAGGGACACATGAAGGGGTGACTACTGGAACGCCAAATTCCTGTTAT 480
Qy 481 GTCCCAACACAGATCAATAGGACGTATCAACCGTGAAATAGCCAAATGTGTACCGACCT 540
Db 481 GTCCCAACACAGATCAATAGGACGTATCAACCGTGAAATAGCCAAATGTGTACCGACCT 540
Qy 541 TCTCATGACAGCGCAACTTATGACTTCAAGTACGGTGTAGAGCTGTACAGGGAGGTGGG 600
Db 541 TCTCATGACAGCGCAACTTATGACTTCAAGTACGGTGTAGAGCTGTACAGGGAGGTGGG 600
Qy 601 AGGTCTCTCGGCGAGAAAACCGTTGGAAGGGGTGCTGCGAGGGGCCCTCCCAAGAAAAT 660
Db 601 AGGTCTCTCGGCGAGAAAACCGTTGGAAGGGGTGCTGCGAGGGGCCCTCCCAAGAAAAT 660
Qy 661 CTTAAGCTCAATGTGGATAGAGATCTTGTGCTGTTTGTTCAAAGTGCATCAGTTGTG 720
Db 661 CTTAAGCTCAATGTGGATAGAGATCTTGTGCTGTTTGTTCAAAGTGCATCAGTTGTG 720
Qy 721 CTCCTCAGAGACGCGGTTGATTATGCTGCTGTAACCTTTGGAACAGATAGAGCAACATC 780
Db 721 CTCCTCAGAGACGCGGTTGATTATGCTGCTGTAACCTTTGGAACAGATAGAGCAACATC 780
Qy 781 GTTAGATGTCTGTATCCAGAGTACGACAGAGAGATGATAGCGCAATCGACAGAGTACGA 840
Db 781 GTTAGATGTCTGTATCCAGAGTACGACAGAGAGATGATAGCGCAATCGACAGAGTACGA 840
Qy 841 GTTCAGGGGATTCGGTGGTGGAGTATCATGCGTGCCTAGAAACGTTCTTCGCGGG 900
Db 841 GTTCAGGGGATTCGGTGGTGGAGTATCATGCGTGCCTAGAAACGTTCTTCGCGGG 900
Qy 901 CTCGGTCTCTGTGTTCGACAAGCTCGAATCCGAACCTGGCAAGCTATGCTTCTATT 960
Db 901 CTCGGTCTCTGTGTTCGACAAGCTCGAATCCGAACCTGGCAAGCTATGCTTCTATT 960
Qy 961 CTGCGAGCAACGGGTTTCAGATTTCGACGCGGATTCGCCGGGACCGACTTGACAGGAAGT 1020
Db 961 CTGCGAGCAACGGGTTTCAGATTTCGACGCGGATTCGCCGGGACCGACTTGACAGGAAGT 1020
Qy 1021 GAGCATATGATGAGTTTATATGATTAAGGCTGGAAGTGTGAGGACGACGACTAATCGC 1080
Db 1021 GAGCATATGATGAGTTTATATGATTAAGGCTGGAAGTGTGAGGACGACGACTAATCGC 1080
Qy 1081 TCGGTTGTGTGACGGGAGGATATCGAATGTTGAGATTGTGCATCTCAAAGTTGCTTTT 1140
Db 1081 TCGGTTGTGTGACGGGAGGATATCGAATGTTGAGATTGTGCATCTCAAAGTTGCTTTT 1140
Qy 1141 AAGCGACACCACTATCGGGGTGAAACAGAACACCGTGTCAAGGGAGCGTCAGAACGTT 1200
Db 1141 AAGCGACACCACTATCGGGGTGAAACAGAACACCGTGTCAAGGGAGCGTCAGAACGTT 1200
Qy 1201 GAGCTTCTAGCAAGAGGGGCAATGACCGATGCTGCGCCCTCGAGCTGTTCTGTGGTG 1260
Db 1201 GAGCTTCTAGCAAGAGGGGCAATGACCGATGCTGCGCCCTCGAGCTGTTCTGTGGTG 1260
Qy 1261 GAATCCATGGCGCGTGTGCTCTCATGACACAGCTGATGGCGACGTCGCTCAGTGGCAG 1320
Db 1261 GAATCCATGGCGCGTGTGCTCTCATGACACAGCTGATGGCGACGTCGCTCAGTGGCAG 1320
Qy 1321 ATGTTCCGCTCAATACCTGCACTTCAAGAACCCAGTTGGCTCTTTCTAGCAGGCGCAGC 1380
Db 1321 ATGTTCCGCTCAATACCTGCACTTCAAGAACCCAGTTGGCTCTTTCTAGCAGGCGCAGC 1380

Qy 1381 ACACCTGATGAGCTCGCGCCAAATTTTATCATTTATCATAGTAATAGTCAAGCGT 1440
Db 1381 ACACCTGATGAGCTCGCGCCAAATTTTATCATTTATCATAGTAATAGTCAAGCGT 1440
Qy 1441 GGCTTGGTTTGTCTCTTTCGACCGTAGTGTGTTTTTTTTTTTTTTTCCCGCAAGTGTGATG 1500
Db 1441 GGCTTGGTTTGTCTCTTTCGACCGTAGTGTGTTTTTTTTTTTTTTTCCCGCAAGTGTGATG 1500
Qy 1501 CGATGAAGTGAATAAGGCACTTGGTTTCCTGTGCATTTGTACACGTTTTCATATATGTAA 1560
Db 1501 CGATGAAGTGAATAAGGCACTTGGTTTCCTGTGCATTTGTACACGTTTTCATATATGTAA 1560
Qy 1561 TCTACTTTCGAGATGATGCAATTTTATAGATGTGGCTGTGGAAGACAAAAA 1620
Db 1561 TCTACTTTCGAGATGATGCAATTTTATAGATGTGGCTGTGGAAGACAAAAA 1620
Qy 1621 AAAAAA 1626
Db 1621 AAAAAA 1626

RESULT 2
AK099850
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) cDNA clone:J013110C10, full insert sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AK099850 1607 bp mRNA linear PLN 24-JUL-2003
Oryza sativa (japonica cultivar-group) cDNA clone:J013110C10, full insert sequence.
AK099850
AK099850.1 GI:32985059
FLI CDNA; CAP trapper.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

1
The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team, Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Otomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Taunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nariikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN: Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Science 301 (5631), 376-379 (2003)
22752273
12869764
2 (bases 1 to 1607)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Inamura, K., Inotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Nariikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,

Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takaku-Akahira, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cdna/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, F., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

PAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Nishikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsumura, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kuribara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Otsu, N., Ota, Y., Saitoh, H., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES

Location/Qualifiers
1..1607
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J013110C10"

ORIGIN

Query Match 53.9%; Score 877.2; DB 8; Length 1607;
Best Local Similarity 80.7%; Pred. No. 3.8e-239;
Matches 1039; Conservative 0; Mismatches 243; Indels 6; Gaps 1;

75 TCAGATGACGACCGTCCCAAGCCACGACGAGTGGCGCACTCACGGGCAACGCTCGCAC 134
DB |||||
55 TCGATGGCGCGCCAAAGTCGTCGAGCCGCTGGCGGCGCTCCCTCCCGCGGCGCG 114
QY |||||
135 CCGCGGATCGCGGCTTCTCGAGTTTCCCGGAGCTCTCTCTCTCTCTCTCTCTCTCTCT 194
DB |||||
115 CCGCGGTTCCGCGGCTTCCCGGAGTCCCGCGGCT-----TCCCTCCGCTTCTCGGT 168
QY |||||
195 GCACCGCTGCCGCTCTGCTCGCTAGAGTGAAGGATCTGGAAACAGCTTTGGAACTA 254
DB |||||
169 CGCGCGCGCGCGCGCTCGCTAGAGTGAAGGCTCTGCAATGTATTTGGAACTA 228
QY |||||
255 CTTTCAGGTGCAACTATGTTGAATCTCATCGGCGGTGGTGTGGTGTGGTGTATCAGTGG 314
DB |||||
229 CTTTCAGGTGCAACTATGTTGAAGTCTCATGGAGCGGTGTGGTGTGGTGTATCAGTGG 288
QY |||||
315 TTGTCCACTAGAAATCCACTCAGTGGCGACACCTACAAAGTTGAACCTGATCGAAGACG 374
DB |||||
289 ATGCCACCCAGAAATCCCACTTACTGAAGCAGATATGCAAGTGAAGTGAAGTGAAGTGA 348

RESULT 3
AR432940
LOCUS

AR432940 1635 bp DNA linear PAT 18-DEC-2003

QY 375 GCCCGGACAGAGCAGAAATAAACCCTCCAAAGGAAGAGACTGATACATGCAAAATTTCTGTC 434
DB 349 GCCAGCCAGAGCAGAAATAAACACCCCAAGAAAGAGACTGACATTTGCCAAATTTCTTTC 408
QY 435 AGGACACATGAAGGGGTGACTACTTGGAAACGCCAAATTTCTTGTATTTGTCCTCCAAACACAGA 494
DB 409 AGGACACATGAAGGAATGACCCTGGGACACCAATTCATGTTTGTCTCCCGAACACAGA 468
QY 495 TCAAATAGGAGTGAATCAACCGTGAATAGCCATGTGTACCGACCTTCTCTATGAGACGC 554
DB 469 TCAGAGAGGGGGTGAATACAGTGAATGGCTTAAGGCTTACAGACCTTACATGTCGATGC 528
QY 555 AACTTATGACTTCAAGTACGGTGTGTAGAGCTGTGTACAGGAGGTGGAGGTCTCTCCGGCAG 614
DB 529 AACTTATGACTTCAATACGGTGTGTAGAGCAGTGCAGGAGGTGGAAGATCATCAGCAAG 588
QY 615 AAAAACCGTTGGAAGGGTGGCTGCAAGGGCCCTCCCAAGAAATAATTTCTTAAGCTCAAAATG 674
DB 589 AGAGACCATTGGAGGGTGGCTGCAAGGAGCTCTTGC AAAAGAAATTTCTTAAGCTCAAAATC 648
QY 675 TGGATTAGAGATCTTGTCTGTTTGTTCCTCAAGTGCATCAGTTGTGCTCCCGAAGACGC 734
DB 649 TGGAGTAGAGATCTTGGCGTTTGTGTCCCAAGGTGCATCAAGTTGTACTACCAAGAGATGC 708
QY 735 GGTTCATTATGGTCTGTAACTTTTGGAAACAGATAGAGCAACATCGTTTAGATGTCCTGA 794
DB 709 CGTTGATTATGACACTGTAACTTTGGAACAGATAGAGCAACATCGTTTAGATGTCCTGA 768
QY 795 TCCAGAGTACGACAGAAAGATGATAGACGCAATCGACAGAGTACAGATTCGAGGGGATTC 854
DB 769 TCCAGATATGCACAGAAAGATGATTGATGCAATCGATAAAGTACGAGTTAGAGGTGATTC 828
QY 855 GGTGGTGGAGTATCAGATCGCTCGCTAGAAAAGTCTCTCGGGGCTCGGTTTCTCTGT 914
DB 829 GATTGGTGGTGGTGGTGCATGCATTTGCAAGAAATGTTCTCTCGTGGGATTTGGCTCTCTGT 888
QY 915 GTTCGACAAAGCTCGAATCCGAATCGGAAAGCTATGCTTTCTATTCTCGGAGCAACGG 974
DB 889 ATTTGACAAACTTGAGGCTGAATTTGGCGAAAGCTATGCTTTCTCTCTGCAAGCAAGG 948
QY 975 GTTCGAGATTGGCAGCGGATTCGCGGGACCGACTTGACAGGAAGTGAAGCAATATGATGA 1034
DB 949 GTTTGAGATCGCAGTGGATTTGCAGGTACTGACTACACTGGAAGTGAAGCAATATGATGA 1008
QY 1035 GTTTTATATGATTAAGGCTGGAAGTGTGAGACACGGACTAATCGCTCGGCTGGTGTGCA 1094
DB 1009 GTTCTTATATGATGAGGCTGGAAATGTGAGAAACAGAACTAATCGTTTCAGGCGGTGCA 1068
QY 1095 GGGAGGGATATCGAATGTTGAGATTGTGCACTTCAAAAGTTCCTTTTAAAGCCGACCATC 1154
DB 1069 GGGAGGGATATCAAAATGTTGAAATATATATCTTCAAGTAGCTTTCAAGCAACCGCGAC 1128
QY 1155 TATCGGGTGAACAGAACACCGTGTCAAGGGAGCGTCAAGAACGTTGAGCTTTCTAGCAAG 1214
DB 1129 TATTGGGAAGAAACAACATACTGTTTCAAGGGAGCATGAGGATGTTGAATTTTAGCAAG 1188
QY 1215 AGGGGCGCATGACCCATGCGCTCGCCCTCGAGCTGTTCTCTGTGTGGATCATGCGCGCG 1274
DB 1189 GGGCGCGCACACCCATGTTGTTCCTTCGCGCTGTGCGGATGGTGGAGTCCATGCGCGCG 1248
QY 1275 GTTGGTCTCATGACACGCTGATGGGCGACGTGCTCAGTCCGAGATGTTTCGGGCTCAA 1334
DB 1249 ATTAGTCTCATGACACCGCTGATGGCGCAATTTGCTCAATGTGAGATGTTTCCACTGAA 1308
QY 1335 TACTGCACTTCAAGAACCCAGTTGGCTCT 1362
DB 1309 CTTTGTCTACAGAACCCAGTTGGCTCT 1336

Db 538 GGAGTGAGATCTGTACAGCGGGTGGTAGATCATCAGCAAGAGAGACTATTGGGAGAGTT 597
Qy 634 GCTCAGAGGGCCCTCCCAAGAAATCTTAAAGCTCAAAATGTGGATTAGAGATCTTTGTCG 693
Db 598 GCTGCTGAGCAGTGTGCTAAAGAAATCTCAAACTTATTTCTGGAGCTGAGGTTCTTGCT 657
Qy 694 TTTGTTTCCAAAGTGCATCAGGTGTGCTCCAGAGAGCGGTTGATTATGGGCTCTGTA 753
Db 658 TATGTTTCTCAAGTTCACCAAGTTGATCTTCTGAGGATTTGATTGATCATCAAGACGTG 717
Qy 754 ACTTTGGAACAGATAGAGAGCAACATCGTTAGATGTCTCTGATCCAGAGTACGCGAGAGAG 813
Db 718 ACTTTAGAGCAGATAGAAAGCAATATTGTCGAGTCCCGGATCCGGATATGCGAGAGAG 777
Qy 814 ATGATAGAGCAATCGACAGAGTACGAGTTTCGAGGGGATTCGGTCGGTGGAGTATCAACA 873
Db 778 ATGATGTGCTGATTGATGCTGTACGAGTGAGAGAGACTCTGTTGGTGGTGTGTTAACT 837
Qy 874 TCGTGTGCTAGAAAGCTTCTCGGGGCTCGGTTCTCTGTGTTTCGACAACTCGAATCC 933
Db 838 TGCAATGTTAGAAATCTCCAGGGGCTTGTGTAACCAGTCTTCGATAAACTTTGAAGCT 897
Qy 934 GAATGGCAAAAGCTATGCTTTCTATCTCGAGCAACGGGTTTCGAGATTGGCAGCGGA 993
Db 898 GAGCTAGCCAAAGCTTGCAATGTCATTACTCGGACAAAGGGTTTGAAGTTGGCAGTGGC 957
Qy 994 TTCGGCGGACCGACTTCAGAGGAAGTGAAGATATGATGATGATTTTATATGATGAAGCT 1053
Db 958 TTTGCGAGCAGCTTCATGACTGGTAGCAGCATAATGATGAATTCATATGATGATGAGCAT 1017
Qy 1054 GGAAGTGTGAGCAGCAGCACTAATCGCTCGGGTGTGTCGAGGGAGGGATATCGAATGTT 1113
Db 1018 GCCGAATCAGAACAAAGCAACAGATCTGTGGAATCCAGGGCGGTATATCAATGGA 1077
Qy 1114 GAGATTGTGCACTTCAAAAGTTGCTTTTAAAGCGCACCACTATCTCGGGGTGAAACAGAAC 1173
Db 1078 GAAGTATCAATATGAGAATAGTTTCAAGCAACTTCAACTATTTCTAGGAAGCAGCA 1137
Qy 1174 ACCGTGTCAAGGAGCGTCAGAACGTTGAGCTTTAGCAAGAGGGCGCCATGACCCATGC 1233
Db 1138 ACTGTGACAAGAGACAAACACAGAACTCATCGCTAGGGGTGCGCATGATCCTTGT 1197
Qy 1234 GTGCGCCCTCGAGCTGTTCTGTGTGGAATCCATGCGCGGTGCTCTCATGAGCAG 1293
Db 1198 GTGCTTCCCGAGCTTCCCAATGTTGAAGCAATGTCGCTGGTGGTCTGTTGATCAG 1257
Qy 1294 CTGATGGCGCAGTGGCTCAGTGCAGATGTTGCGGCTCAATATCTGCACTTCAAGAACC 1352
Db 1258 TTATGGCTCAGTATTCACAGTGTATGATGTTCCCAATCAATCTTGAACCTACAGNACC 1316

RESULT 6
LOCUS LECHOSYNB 1802 bp mRNA linear PLN 06-FEB-1996
DEFINITION L.esculentum chorismate synthase 2 precursor.
ACCESSION Z21791
VERSION Z21791.1 GI:410483
KEYWORDS chorismate synthase 2.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 1802)
AUTHORS Gorlach,J., Schmid,J. and Amrhein,N.
TITLE Differential expression of tomato (Lycopersicon esculentum L.) genes encoding shikimate pathway isoenzymes. II. Chorismate synthase
JOURNAL Plant Mol. Biol. 23 (4), 707-716 (1993)
MEDLINE 94072719
PUBMED 8251624
REFERENCE 2 (bases 1 to 1802)

AUTHORS Gorlach,J., Raesecke,H.R., Abel,G., Wehrli,R., Amrhein,N. and Schmid,J.
TITLE Organ-specific differences in the ratio of alternatively spliced chorismate synthase (LeCS2) transcripts in tomato
JOURNAL Plant J. 8 (3), 451-456 (1995)
MEDLINE 96045551
PUBMED 7550381
REFERENCE 3 (bases 1 to 1802)
AUTHORS Goerlach,J.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-1993) Joern Goerlach, Institute of Plant Sciences, Swiss Federal, Institute of Technology, Sonneggetr. 5, Zuerich, CH-8092, Switzerland
FEATURES
source Location/Qualifiers
1..1802 /organism="Lycopersicon esculentum"
/mol_type="mRNA"
/strain="UC82B"
/db_xref="taxon:4081"
/clone="cLeCS2"
42..1337
/EC_number="4.6.1.4"
/codon_start=1
/product="chorismate synthase 2"
/protein_id="CAA79854.1"
/db_xref="GI:410484"
/db_xref="GOA:Q42885"
/db_xref="UniProt/Swiss-Prot:Q42885"
/translation="MASSMLTKQFLGAPFSGGQOPSKLCSNLRFPFTHRSQPKRL EIQAGNTFNGYFRVTTFGSHGGVCIIDGCPRLPLSESDMQVELDRRFGQSR ITPRKETDTCKISSGTADGLTTGSPIKVEVPNTDQRGNDYSEMSLAYRPSHADATYDF KYGVSQGGSSARETIGRVAAGAVAKILKLYSGTEILAVVSQVHVVLPELDLVD NQIVTLEQIESNIVRCNPENPEYAEKMGADYVVRGDSVGVVVTCTIVNVRGLGTFV FDKLEALAKMSLPATKGFEPFGFAGTFMGSEHDEFNDEHDIQIKTKNRSQG IQGGISNGEIIKNNRFAKPTSTIARKOHTVSRDKHETELIARGHDPCCVVPRAVPMVE AMVALVLVDQLMTQYACMLFPVNLTLQEPLOPSTTKSA"
42..185
sig_peptide
186..1334
mat_peptide
/product="chorismate synthase 2"
/EC_number="4.6.1.4"
1802
polyA_site
ORIGIN
Query Match 35.3%; Score 574.6; DB 8; Length 1802;
Best Local Similarity 67.9%; Pred. No. 1.5e-152;
Matches 802; Conservative 0; Mismatches 379; Indels 0; Gaps 0;
Qy 172 TCCTCTCTCCCTCGCCTTGCCTGTCACCGCTGCGGCACTGCTCGCCTAGAGGTGAAGCA 231
Db 126 TCTTCAAACTCCGATTCCCTACTCACCGATCACAGCCAAACGACTAGAGATACAAGCT 185
Qy 232 TCTGGAACAAGTTTGGAACTACTTTCAGTTGCAACCTATGGTGAATCTCATGGGGGT 291
Db 186 GCTGGTAATACATTTGGAAATATTTCCGGCTTACACACGTTTGGAGAATCTCAGGTGGT 245
Qy 292 GGTGTTGTTGTGTTATCAGTGGTTGTCACCTAGAAATCCACTCAGTGAAGCAGACTA 351
Db 246 GGTGTCGGTTGTATTAATTGATGATGTCCTCCCTCGACTCCACTCTCAGAGTCGATG 305
Qy 352 CAAGTTGAATCGCATCGAAGACGGCCCGGACAGAGCAGAATAACCTCCAAGGAAGAG 411
Db 306 CAAGTGAACCTTGACAGGAGGCGCCAGGTCAAAGCCGAATTACCACACCAAGAAAAGAG 365
Qy 412 ACTGATCATGCAAAATTCGTGAGGACACATGAAGGGGTGACTACTTGGAAACCCCAATT 471
Db 366 ACTGATCTTGCAAAATATCATCAGGCACCTGCTGATGTTTGACTACGGGATCTCCCAATC 425
Qy 472 CTGTTATTGTCCTCCAAACACAGATCAATATAGGCACTGATCACCGTGAATAGCCCAATGTG 531
Db 426 AAAGTTGAAGTACCAACACTGACAGAGAGAAATGACTATAGTGAATGTCCTCTTGTCT 485
Qy 532 TACCGACCTTCTCATGCGAGACGCAACTTATGACTTCAAGTACCGTGTAGAGCTGTACAG 591


```
Db 804 GTCTCTCAAGCCCAAGAGTTGTTCTTCCAGAGGGTTGGTTGATCATGAGACTATCA 863
Qy 757 TTGGAACAGATAGAGAGCAACATGTTAGATGTCCTGATCCAGAGTACGCGAGAGAAGATG 816
Db 864 CTCGAACAGATTGAAGTAACATCTTAGTGTGCCAGATTCAGAAATACGCCGAGAGATG 923
Qy 817 ATAGAGCAATCGACAGAGTAGAGTTCGAGGGGATTCGTCGGTGGAGTATCACATGC 876
Db 924 ATTGCTGCTATTGATGCTGTCCGAGTGAAGGAGACTCTGTGGTGGAGTTGTGACATGT 983
Qy 877 GTCGCTAGAAAGCTTCTCGCGGGCTCGGTTCTCTGTTGTCGACAACTCGAATCCGAA 936
Db 984 ATTATGAGGAAGCTCCAGTGGGCTTGGTCCACAGTATTGTATTAACATTGAGGCTGAG 1043
Qy 937 CTGGCAAAAGCTATGCTTTCTATCTCGAGCAACCGGTTGAGATGGCAGCGGATTC 996
Db 1044 CTTGCCAAAGCGTGATGTCATTACCAACCAAGGGGTTTGAATTTGGGAGTGGATT 1103
Qy 997 GCGGGACCGACTTGCACAGGAGTGAAGCATATGATGATGTTTATATGATAGGCTGGA 1056
Db 1104 TCAGGCACATTTCTGACTGGTAGTGAACATATGATGATGTTCTACACAGACGAGATGGA 1163
Qy 1057 AGTGTGAGGACACGACTAATCGCTCGGGTGGTGTGACGGGAGGATATCGAATGTTGAG 1116
Db 1164 AGAATCAGGACTAGAACAAATCGATCTGGTGGATACAGGGTGGTATATCCACGGAGAA 1223
Qy 1117 ATTTGCTACTTCAAAAGTTGCTTTTAAAGCCGACACCATCTATCGGGGTGAACACAGAACCC 1176
Db 1224 ATCATAAACATGAGAATAGCTTTCAAGCCGACATCTACAATTTGGAAGAAAGCAGAACT 1283
Qy 1177 GTGTCAAGGAGCGTCAGAACGTTGAGCTTCTAGCAAGAGGCGCCATGACCCATGGCTC 1236
Db 1284 GTGACGAGAGAACGTTGAAGAGATTGAATCTATAGCAGTGTGCGCCACGACCTTGTGTT 1343
Qy 1237 GCGCTCTGAGCTGTTCTCTGTGTGGAATCCATGCGCGGTTGGTCTCTATGACCGAGCTG 1296
Db 1344 GTACTCTGAGCTGTTCCAAATGGTGGAGCTATGGTAGCCCTGGTCTCTGATCACTG 1403
Qy 1297 ATGGCGCAGTGGCTCAGTGCAGATGTTTCGCGCTCAATATGCACTTCAAGAACCC 1352
Db 1404 ATGCTTCAACATGCTCAGGGCAACCTTTTCTCCATCAATCTGCTTCAAGAACCC 1459
```

RESULT 8

LOCUS CQ805346 1311 bp DNA linear PAT 10-MAY-2004
DEFINITION Sequence 1757 from Patent WO2004035798.
ACCESSION CQ805346
VERSION CQ805346.1 GI:47111307

KEYWORDS

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.

REFERENCE

1 Inze, D., de Veylder, L. and Vlieghe, K.

Identification of novel e2f target genes and use thereof

Patent: WO 2004035798-A 1757 29-APR-2004;

CropDesign N.V. (BE)

FEATURES

source

1..1311

/organism="Arabidopsis thaliana"

/mol_type="unassigned DNA"

/db_xref="taxon:3702"

ORIGIN

Query Match 32.9%; Score 535.4; DB 6; Length 1311;
Best Local Similarity 66.9%; Pred. No. 2.3e-141;
Matches 761; Conservative 0; Mismatches 376; Indels 0; Gaps 0;
Qy 216 CCTAGAGGTGAAGGCATCTGGAACACGTTTGAAACTACTTTTCAGGTTGCAACCTATGG 275

```
Db 135 CTTCCAGATACAGCTACTCGAAGTTCAATGGGACTCATTTTCGAGTTTCAACTTTGG 194
Qy 276 TGAATCTCATGGGGTGGTGTGTTGTGTATTCAGTGGTTGCCACCTAGAAATTCACCT 335
Db 195 AGAATCACATGGAGGAGGAGTGTGTTGTATCAATTGATGGTTGCTCCTCGTATTCACCT 254
Qy 336 CACTGAGGCGACCTACAGTTGNACTCGATCGAAGACGCCCGGACAGAGAGCAATAAC 395
Db 255 TACTGAACTCTGATTTGCAATTCGATCTCGATAGAGGAGCCCTGGTCAGAGAGGATCAC 314
Qy 396 CTCCACAAGCAAGAGAGACTGATACATGCAAAATTTCTGTCAAGGACACATCAAGGGGTGAC 455
Db 315 AACTCTTAGAAGAGAGACTGATACTTGGCCGGATATCGTCTGGAGTCTCTGAGGAATGAC 374
Qy 456 TACTGGAACCCCAATCTTGTATTTGTTCCAAACACAGATCAAAATAGGAGTATGATCCG 515
Db 375 GACAGGAACACCTATCCATGTGTTGTACCAACACAGATCAGAGAGGACTTGTATTACAG 434
Qy 516 TGAATAGCCCAATGTGTACCGACTTCTCATGACAGCGCAACTTATGACTTCAAGTACGG 575
Db 435 TGAATGTGCGGTTGCCCTATAGACCATCGCATGCTGATGCAACTTATGACATGAAGTATGG 494
Qy 576 TGTTAGAGCTGTACAGGGAGTGGGAGGTCTCGGGCAGAAACCCGTTCGAAAGGGTGGC 635
Db 495 TGTCAGATCAGTGCAGGGTGGAGGAGATCTTCAGCTAGAGAGACCATTGGAGAGTTGC 554
Qy 636 TGCAAGGGCCCTCCCCAAGAAAAATTTCTTAAGCTCAAAATGGGATTAGAGATCTTGTGCT 695
Db 555 TCCTGGAGCTTTGGCCAAAGAAAAATTTTGAAGCAATTTGACAGGAACCTGAGATCTTCGCTA 614
Qy 696 TGTTTCCAAAGTGCATCAGGTTGTCTCCAGAGAGCGCGTTGATTTATGGTCTGTAAAC 755
Db 615 TGTCTCGCAAGTTTCAACATGTGTTACTTCCAGAAAGAAATTTGTAGACCAACAGAAATTAAC 674
Qy 756 TTTGGAACAGATCAGAGCAACATCGTTAGATGTCCTGATCCAGAGTACGCGAGAGAGAT 815
Db 675 ACTGACACAGATAGAAATAACATTTGTCAGATGCCCTAATCCCGAGTATCGGGAAGAT 734
Qy 816 GATAGACGCAATCGACAGATGACGAGTTTCGAGGGGATTCGGTTCGGTGGAGTATGATC 875
Db 735 GATAGCTGCGATTTGATGCTGTGAGGACAAAGAGGAACTCTGTTGGTGGTGTGTTGACCTG 794
Qy 876 CGTCGCTAGAAACGTTCTCTCGCGGCTCGGTTCTCTGTTGTCGACAGCTCCGAATCCGA 935
Db 795 CATTTGTTGGAATGCTCCAGTGGGCTTGGTACACCGGTTTCGATAAATCTTGAAGCAGA 854
Qy 936 ACTGGCAAAAAGCTATGCTTTCTATTCCTGCGAGCAACGGGTTTCGAGATTCGACGCGAT 995
Db 855 ACTGGCAAAAAGCTTGTATGTGCTACCTGCAACTAAGGGATTTGAGTTTGGNAGCGGCT 914
Qy 996 CGCCGGACCGACTTGAAGAGTGAAGATATGATGATGATGATTTATATGATAGGCTGG 1055
Db 915 TGCAGGTACCTTTTGTAGCTGGTCTTGAACACAATGATGAGTCTTATACCGATGAAATGG 974
Qy 1056 AAGTGTGACAGACCGGACTAATCGCTCGGTTGGTGTGACGGGGGATATCGATGTTGA 1115
Db 975 AAGAATACGTAACCAAGAACCAACCGATCTGTGGAATTCAGGGAGGGATCTCAATGTTGA 1034
Qy 1116 GATTGTGCACTTCAAAAGTGTCTTTTAAGCCGACACCATCTATCGGGGTGAAACAGAAAC 1175
Db 1035 AATAATAAATCATGAGAGTAGCTTCAAGCCCAACATCAACAAATTTGAAGGAAGCAAAATAC 1094
Qy 1176 CGTGTCAAGGGAGCGTCAGAAAGTGTGAGCTTTTAGCAAGAGGGGCCCATGACCATCGGT 1235
Db 1095 GGTAAACAGAGACAAGGTTAGAAACCCGAAATGATGTCGCGTGTGTCATGATCTCTGTGT 1154
Qy 1236 CGCCCTCTGAGCTGTTCTCTGTGTGGAATCCATGCGCGCTGTGCTCTCATGGACAGCT 1295
Db 1155 TGTTCACAGAGCTGTGCAATGGTGAAGCAATGTTGGTCTCTAGTTCTTGTGGATCAATT 1214
Qy 1296 GATGGCAGCTGCTCAGTGCAGAGATGTTTCGCGCTCAATACTGCACTTCAAGAACCC 1352
```

Db 1215 GATGGCCCAATACGCAATGCCATTGTTTCCATAAATCCAGAGTTCAGGAACC 1271

RESULT 9

BT015785

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BT015785

Arabidopsis thaliana

1311 bp mRNA linear

AT1g48850 mRNA, complete cds.

BT015785

BT015785.1

FLI CDNA.

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1311)

Shinn, P., Chen, H., Cheuk, R., Kim, C.J. and Ecker, J.R.

Arabidopsis ORF clones

Unpublished

2 (bases 1 to 1311)

Shinn, P., Chen, H., Cheuk, R., Kim, C.J. and Ecker, J.R.

Direct Submission

Submitted (02-OCT-2004)

(SIGNAL), Plant Biology Laboratory, The Salk Institute for

Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,

USA

FEATURES

source

1..1311

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="1"

/clone="U67693"

/ecotype="Columbia"

/note="This clone is in pUNI 51"

1..1311

/note="unknown protein"

/codon_start=1

/product="At1g48850"

/protein_id="AA090075.1"

/db_xref="GI:53749180"

/translation="MASSSLTSKLSILGSLPRLSLSPAVQISLRTQTRK

NFQIQTSSYTHFRVSTFGSHGGVCIIDGCPRIPLTESDLPDLDRRRQCS

RTTPKRETDTCRISSGVSEGMTTGPPIHVFPNTDQRLDYSEMSVAVRSHADATY

DMKYVRSVQGGRSARETICRVAPGALAKKILKQFAGTEILATVQVHVLPPEEL

VQHNLLEQIENNVRCNPBYAEKMTAAIDAVRTKNSVGVVTCIVRNAPRLGRT

PVFDLEALAKASLPATGFFGSGFAGTFLTGLRHNDFFYDENGRIITRNS

GGIQGINSGETIINRVAFKPTSTIGRKQNTVTRDKVETEMIAIRGRHDPVVPRAVPM

VEAMVALVLDQLMAQYAQCHLFPINPELOEPLQIEQPONATAL"

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

0;

Mismatches

376;

Indels

0;

Gaps

0;

Length

1311;

275

194

335

254

395

314

455

374

515

Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayaishizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
Arabidopsis cDNA clones
Unpublished

TITLE
JOURNAL
REFERENCE
AUTHORS

2 (bases 1 to 1603)

Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayaishizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.

TITLE
JOURNAL

Direct Submission
Submitted (18-SEP-2001) Salk Institute Genomic Analysis Laboratory (SIGnAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayaishizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs (Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Shinn,P., Banh,J., Bowser,L., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Davis,R.W., Theologis,A., and Ecker,J.R.

Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

source

1..1603
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAFL09-47-O14 (R14512)"
/ecotype="Columbia"

5'UTR

CDS

123..1433

/note="putative chorismate synthase"
/codon_start=1
/product="At1g48850/T24P22_3"
/protein_id="RA109759.1"
/db_xref="GI:1598284"
/translation="MASSSLTSKSLGSKLSSLSBELRRLSSPAVQISLRQTRK
NFQQTSSGTHFRVSTFGESGGGVGCIIDGCPRIPLTESDLDQDRLRRPGQS
RITPRKETDCRISVSEGMGTPIHVFPNTDQRLDYSMSVYRPSHADATY
DMKTVRSVQGGRRSARETIGRVAPALAKILKQFAGTEILAYVSVHVVLPPEL
VDHENLLEQENNTVRCNPPEYAEKMIADAVRTKGNVSGVVVTCVIRNAPRLGT
PVFDKLEAKACMLPATKPFEGSGFAGTELTGLEHNDPEFYTDENGRIRRTNRS
GGIQQGINSGLIINRVAPKPTSTIGRKQNTVTRDKVETEMIAGRHDPVVRPAPVM
VEAMVLVLDLMAQYAQCHLFPINPELQBPQLQEQPNATAL"
1434..1603

3'UTR

ORIGIN

Query Match 32.9%; Score 535.4; DB 8; Length 1603;

Best Local Similarity 66.9%; Pred. No. 2.4e-141;

Matches 761; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

QY

216 CCTAGAGGTGAGGCATCTGGAAACACCGTTTGGAAACTACTTTTCAGGTGGCAACCTATGG 275

Db 257 CTTCCAGATCAAGCTACTGGAAGTTTCATATGGGACTCATATTTTCGAGTTCCTCAACTTTGG 316
QY 276 TGAATCTCATGGGGTGGTGTGGTTGTGTATCAGTGGTTGTCCACCTAGAATTCACACT 335
Db 317 AGAATCAATGAGGAGGAGTGGTTGTGTATCATTTGATGGTTGTCTCTCCTCGTATTCACACT 376
QY 336 CACTGAGGCAGACCTACAAGTTCGAATCGATCGAAGACGCCCGGACAGAGCAGAAATTAAC 395
Db 377 TACTGAATCTGATTTGCAATTCGATCTCGATAGAGGAGGCGCTGGTCAGAGCAGGATCAC 436
QY 396 CTCCACAAGGAAGAGAGACTGATACATGCAAAAATTTCTGTCAAGGCACACATGAAGGGGTGAC 455
Db 437 AACTCTTAGAAAAGAGAGACTGATCTTGGCGGATATCGTCTGGAGTCTCTCGAAGGAATGAC 496
QY 456 TACTGGAACGCCAATTTCTGTATTTGTTGCCAAAACACAGATCAAAATAGGAGTGTATCACCG 515
Db 497 GACAGGAACCACTATCCATGTGTGTATGTTGATCCAAAACACAGATCAGAGAGGACTTGTATTACAG 556
QY 516 TGAATAGCCAATGTGTACCGACCTTCTCATGAGACGCAACTTATGACTTTCAAGTACGG 575
Db 557 TGAATGTCTGGTTCCTATAGACCATCGCATGCTGATGCAACTTATGACATGAAGTATGG 616
QY 576 TGTTAGAGCTGTACAGGAGGTGGAGGTCTCTCGGCAGAAAAACCGTTGGAAGGGTGGC 635
Db 617 TGTGAGATCAGTGCAGGGTGGAGGAAGATCTTCAGCTAGAGACCACTTGGAGAGTTGC 676
QY 636 TGCAGGGGCCCTCCCAAGAAAAATTTCTTAAGCTCAAAATGTGGATTAGAGATCTTGTGCTT 695
Db 677 TCTTGGAGCTTGGCCCAAGAAAAATTTTGAAGCAATTTTGCAGGGAATCTGAGATTTCTTGCTTA 736
QY 696 TGTTCCTCAAGTGCATCAGGTGTGCTCCAGAAAGACGCGTTGATTTATGGTCTGTAAAC 755
Db 737 TGTCTCGAAGTTTCACTGTTGTACTTCCAGAAAGATTTGGTAGACCAAGAGATTTTAAAC 796
QY 756 TTTTGGAAACAGATAGAGAGCAACATCGTTAGATGTCTCTGATCCAGAGTACGACAGAAAGAT 815
Db 797 ACTCGAAACAGATAGAAAAATTAACATTTGTGATGCTCCCTAATCCCGAGTATCGGAAAAAGAT 856
QY 816 GATAGAGCAATCGACAGAGTACAGGTTCAGAGGGATTCGGTGGTGGAGTATCATCATG 875
Db 857 GATAGCTGCGATTTGATGCTGTCTAGGACAAAAGGGAACCTCTGTTGGTGGTGTGTGACCTG 916
QY 876 CGTCTGTAGAAAACTTCTCTCGCGGCTCGGTCTCTCTGTGTTCGACAAGCTCGAATCCGA 935
Db 917 CATTTGTCGAATGCTCCAGTGGCTTGGTACACCGTTTTCGATATAACTTGNAGCAGA 976
QY 936 ACTGGCAAAAGCTATGCTTTTCTTCTCTGCGACCAACGGTTTCGAGATTCGACGCGGATT 995
Db 977 ACTGGCAAAAGCTTGTATGTCTGCTACCTGCAACTAAGGGATTTGAGTTTGGAGCGGCTT 1036
QY 996 CGCGGGACCGACTTGAACAGAGTGAACATATGATGAGTTTATATGATAGGCTGG 1055
Db 1037 TGCAGGTACCTTTTGTGACTGCTTGTGAACAAATGATGAGTCTTATACCGATGAAATGG 1096
QY 1056 AAGTGTGAGGACCGGACTAATCGCTCGGTGGTGTGCGAGGAGGATATCGAATGTGA 1115
Db 1097 AAGNATAGTACAGAACCAACCGATCTGTTGAAATTCAGGAGGAGTCTCAATGTGTA 1156
QY 1116 GATTGTGACTTCAAGTTTCTTTTAAAGCCGACACCATCTATCGGGGTGAAACAGAACAC 1175
Db 1157 AATAATAAACATGAGAGTAGCTTCAAGCCAAACATCAACAAATTTGGAAGGAAGCAAAATAC 1216
QY 1176 CGTGTCAAGGAGGCGTCAGAACGTTTGAAGCTTCTAGCAAGAGGGGCGCATGCCATCGCT 1235
Db 1217 GGTAAACAGAGACAAGGTAGAAACCGGAAATGATTTGGCGGTGGTGGTCTGATGATCTCTGTGT 1276
QY 1236 CGCCCTCTGAGCTGTTCTCTGTGTGGAATCCATGGCCGCGTGTGCTCTCATGGACCAAGCT 1295
Db 1277 TGTTCACAGAGCTGTGCCAATGTTGGAGCAATGGTGGCTCTAGTCTTCTTGGATCAATT 1336
QY 1296 GATGGCGCAGTGGCTCAGTGCAGAGATGTTTGGCGCTCAATATCTGCACTTCAAGAAC 1352
Db 1337 GATGGCGCAATAGCGCAATATGCTTGTGTTTCCAAATAATCCAGAGTTGCAGGAAC 1393

[illegible]

source	1. .339650 /organism="Nostoc sp. PCC 7120" /mol_type="genomic DNA" /db_xref="taxon:103690" /note="synonym:Anabaena sp. PCC7120" complement(98. .2698) /gene="all0594" complement(98. .2698) /gene="all0594" /note="ORF ID:all0594" /codon_start=1 /transl_table=11 /product="aminopeptidase" /protein_id="BAB72552.1" /db_xref="GI:17129940"
gene	/translation="MPHSYSFDODSNAHKSPELPGARPHYNDRPGQVHEIFLDNLNLD IFQSYEGSCSIRLPIDIRNGIDRLTLDAVNLQIRSVQVDEVGDDEYDGEQVLIQLSQ PEIRGRILIIAYAVEKPGRIYFIQPSDHPYKTVQVMTQGEDEDSFNPFCFDYP GQLSTSEIRVRYPKSLMAISNGELIDTQDGDGHKIYHMSQQQIHPTYLMTLAVADF IRDEWRGKPTVYVKGREADAKRSMGKTPRMIEFLSEKYGAYAPPKYAOVCVSDFI FGCMGNTSTLLTDRCLDERAALDNRTESLVVHELAWHGFGLLVIKHWSHAWIKE GNASYSEVMWTEQESAEAAVYRLLEARSYLSSESGRYRRPMWTHVYREALIYDRH TYEKSGCYHMRALHGEDLPWQAIQTIVRDNAHKIVETIDLLRAIEKATGNLAPLF DQYVRGHPDPKAYSWDGEAKLAKVTVTQVQPADKDKOVDFDLRIPIGFQYNQTEA AAKLQTFVVRHERQSFYFPLETERPDFISFDVGNHMLKTVSLSEYPIPELKAQLASDP DPSIRYMAAALAKKGLLEATLAASLKNBPFVGVRVEVAKOLAEIKLDOAFDGLVE GLADNAVYRRVAQALAKIKTLASVKAQVGVKQDASVYVESAAACLAGAIAINF DDKSKEKAIKLLQSVLERAGNEVRSAGLGLAELKTSALDILLLEYTKLGVPQ PURLAALQADGKISQGVSPVNLRIIARIGELSKENFFLTLMAVVAALGQMETPKANG VLRSLADQTAGDRFRFAEEELANVQKNIGTDNALRKLREEDQLKQOQNQLSRLEN LEAKSQTKVQ"
gene	complement(2862. .3059) /gene="ael0595"
CDS	complement(2862. .3059) /gene="ael0595" /note="ORF ID:ael0595 similar to cytochrome d oxidase chain I" /codon_start=1 /transl_table=11 /protein_id="BAB72553.1" /db_xref="GI:17129941" /translation="WRHCVHPILLSLPLYKIIILNIQWKPSDRPHMVGLIYAPRVN IAIGPFLAGLMVLSTLFEVR" complement(3151. .3810) /gene="all0596" complement(3151. .3810) /gene="all0596" /note="ORF ID:all0596 unknown protein" /codon_start=1 /transl_table=11 /protein_id="BAB72554.1" /db_xref="GI:17129942" /translation="MGVCHSKDVSBFCLEGRPFMEFVIKDGYKLKGLLLLTADGECYVK LSKQLRFTFWRLPPTGLIQVVGKKYDAKTKVTLKABAVMAARTETSSAVKQVPM ENTKPKPKAKATILVCKQSCDMKRGKALCOALEATISDRGLEQDVITIKGTGCMKNC KAGPNLVNPDKTRITRIQADQVPLMKNHFDADSQQSDSHQAVISNTNKFEPFRVSA F"
gene	complement(3877. .4134) /gene="ael0597"
CDS	complement(3877. .4134) /gene="ael0597" /note="ORF ID:ael0597 hypothetical protein" /codon_start=1 /transl_table=11 /protein_id="BAB72555.1" /db_xref="GI:17129943" /translation="NKPQSSVEEKSVEVVDWDRWQVYQRLQLELIPWCNEANQPLK VELVTPMTVQWQSVMSVQRTSSRQDLIYSLEINWOSRYPYS" 4451. .5146 /gene="alr0598" 4451. .5146
gene	/gene="alr0598" /note="ORF ID:alr0598 unknown protein" /codon_start=1 /transl_table=11 /protein_id="BAB72556.1" /db_xref="GI:17129944" /translation="WVTRMPLISTLSLVLLQCPVIVQAGSENTGDVASTALQOV IAANLMTNSADGNFYPERLMNRABELASILVKAFLNKQEAANQPKSITVADVPRSHWA YQDIQTVLKTDVMKGYRNLFPNQRVTRAEALAIFAQAYGVQFQSDVDVNEILAPYP DANSIPNWARAIATVITEGPIINTDAQNNTPLRPMTRGDMAYLLSKYLQKQKQPD PVPTTDSQPLP"
gene	5596. .7503 /gene="alr0599"
CDS	5596. .7503 /gene="alr0599" /note="ORF ID:alr0599" /codon_start=1 /transl_table=11 /product="1-deoxy-xylulose 5-phosphate synthase" /protein_id="BAB72557.1" /db_xref="GI:17129945" /translation="MHLSEITHPNQLHGLSIRLQOIQAIRQDKHLQTVABFGGHLGP GLGVVELTGLYQTLDDLRDKVIMDVGHQAYEPHKLITGRYSDFHTLRQKDGTAQYVKR GENKFDHFGAGHASTISAALMALARDMNGEKKAVAVIGDALTGMALEAINHAG HLPTKLLVLNDNEMSTSPNVGAI PRVYNKWRSLSPVQFIKDNFEQFPHI PVGES LSPELGRICEGKRLAVPKVCAVFEELGFTTNGPVDGHNLEELIATFOQAHQIAGPVL VHVATIKGKGYELAEKQOVGHAQTTPNLTITTKAI PSNKPKPPAYAKVFSHTLVKLA QNPKIIGITAMATGTGLDKLQAKLPNQYIDVIAEQHAHVTLAAGLATTEGMRPVAAIY STFIQRAYDQIIHDVQIQNLFPVFCLDRAQIVGSDGPTHQGMQDIAYLRCPNIVMA PKDAEMQRMVVVTGIEHTTGPIMRFPNGYGVPLMEGWEPLIEIGKIELRNGDDV LIIGYQVTPVSMQAAETLSHGEIATVINA RFVKPLDTELI VPLAKIGRVTLEEG CVMGSGSAVABALDDVVVPVKRIGIPDVLVVEHATPDESKAELGLTSRQIAERVLQ AYFQKQSPVS"
gene	7793. .8281 /gene="alr0600"
CDS	7793. .8281 /gene="alr0600" /note="ORF ID:alr0600 unknown protein" /codon_start=1 /transl_table=11 /protein_id="BAB72558.1" /db_xref="GI:17129946" /translation="MADELKVKIECFILSKLHMKLNFMMKKLSNLGRWTATTVFCL SAIAFVWGAFFADTSAMDNSTYLIATSDATSLRNKSKFIRIDKNVKAARTNA SRVDDATNGSTAEKANKADASRIQKGAEDARTETAVDNLNNAVERTIDNIKGAFT NN" complement(8399. .9469) /gene="all0601" complement(8399. .9469) /gene="all0601" /note="ORF ID:all0601 hypothetical protein" /codon_start=1 /transl_table=11 /protein_id="BAB72559.1" /db_xref="GI:17129947" /translation="MSIVFRDLLKKVSGNHTGESLNRVEAATATKMFGLGEATPAQI GAFLIAHRIKRPTEGELGMDLDAFDELGPKLOPISSKRPVIVFGIPYDGRTRTAPIS VTALLMATVQPVVHGGDRLPKYGLPLTEIMQGLGVNWTLSLEKTOEVFAQTIG FVYTPKHPLTNSIWEYEDQLGKRPPLATMELIWCYPAGDAHVLAGYVHPPEGMQV ALELRGVSKTILVKLGSGCDLPRDRTAISLSSASELERVFLSTHYEGTTKNVPL GSTEELLTIQDVLGKSCSELQQTALWNGGFYLWRSGICADMRSGIAKAEELNSGLV AAKLQELIQVNVSTENLSIT" complement(9680. .10618) /gene="ntcB" complement(9680. .10618) /gene="ntcB" /note="ORF ID:all0602 ntcB gene product" /codon_start=1 /transl_table=11

```
/product="nitrogen assimilation transcriptional activator"
/protein_id="BAB72560.1"
/db_xref="GI:17129948"
/translation="MRLEQLAFLAIAIATBGSFQKAAKCGVQSTISRQIQSLDADLG
LELFRHTNHSKLITGELRLPRARKICLEWETATQELGDLIAGKQELCIAAAIHSLGQ
SYLPVLOKFCRDYPEVOLRVYSLGSDRALVKDGLVDLAI VNNRFLTTGRMDVVE
VLYDEPIELLTAAHPLAAVERVPWSLVRYPQVFKDGYGMORLAVQEKPERLEATLQ
AAWLEPITADRGVVRQSELIALLPSSALVEARLDPTLAVRPLANSALPENSGLTRRV
NMVTTQDLRLPPIKHFQWLVRENIPIIVERQSRAS"
/gene
CDS
complement(10861..11166)
/gene="all10603"
complement(10861..11166)
/gene="all10603"
/note="ORF_ID:all10603
putative transcriptional regulator"
/codon_start=1
/transl_table=11
/protein_id="BAB72561.1"
/db_xref="GI:17129949"
/translation="WKTKPFDLKKMTPEQQAASEMQANLVLLKLILSELRESLGQT
QGDVAKNMGVVQSALSKEIBHDDDIQISTLSRYIKSLGSLTIIACFPDQEVVISQFD"
complement(11163..11522)
/gene="all10604"
complement(11163..11522)
/gene="all10604"

Query Match 21.6%; Score 351.4; DB 1; Length 339650;
Beat Local Similarity 57.9%; Pred. No. 3.4e-88;
Matches 647; Conservative 0; Mismatches 461; Indels 9; Gaps 1;

QY 209 CTGCTCGCTAGAGGTGAAGGCATCTGGAAACACAGTTTGGAAACTACTTTCAGGTTGCAA 268
DB 232378 CGCGTCTCTTTTCTCAAGATCATGGGCATACTTTTGGACATCTGTTTCGGATCACTA 232319

QY 269 CCTATGTTGAATCTCATGGGGTGTGTGTGTGTATCATGTTGTTGTCTCCACCTAGAA 328
DB 232318 CTTTGTGGCGAGTCACACGGTGGAGCGTGGGGTGTGATTGACGGCTGTCTCGCTAC 232259

QY 329 TTCCACTCAGTGAGCGACCTACAACTTGAATCGATCGAAGACGGCCCGGACAGCA 388
DB 232258 TCGAAATTTCTCTGAGGAAATTCACCTGGAGTTAGACCGCAGACGCCCGGACAGATA 232199

QY 389 GAATAACCTCCACAAGGAGGAGCTGATACATGCAAAATTTCTGTCAGGACACATGAAG 448
DB 232198 AGATTACACCGCTCGTAAGAGCAGACACTGTGAGATTTGTCTGGGTATATGAAG 232139

QY 449 GGGTGACTACTGGAAACGCCAATTTCTTTGTTATGTTCCCAACACAGATCAAAATAGCAGTG 508
DB 232138 GCAAAACCTCTGGAAACGCCAATATCAATTTGGTGGAAATAAGAACTACTCGTCCCAAG 232079

QY 509 ATCACCCTGTAATAGCCAAATGTATACCGACCTTCTCATGTCAGACGCAACTTATGACTTCA 568
DB 232078 ATTATGATGAGATGGCGCAAAAGTATCGTCTTCCCATGCTGATGCAACTTATGATGCTA 232019

QY 569 AGTACGGTTGTAGACTGTACAGGAGGTGGAGGTCTCGGGCAGAAAACCGTTGGAA 628
DB 232018 AATATGGGATTCGTAATTTGGCAAGGTGGCGGTAGATCTCTAGCGGGGAGACAATTTGTA 231959

QY 629 GGGTGGCTGCAGGGGCCCTCCCAAGAAAATTTCTTAAGCTCAAAATGTGGATTAGAGATCT 688
DB 231958 GAGTAGCCGCGAGTGGATCGCTAAATAAATTTCTCCGGC-----AAGTTGCTAATG 231908

QY 689 TGTCTGTTGTTTCCAAAGTGCAATAGGTTGTGCTCCAGAGACGCGGTTGATATGGGT 748
DB 231907 TAGAAGTTATAGGCTACGTCAAACGCTATCAAGACTTGGAAAGGTGTTGTAGATCCCAATA 231848

QY 749 CTGTAACTTTTGAACAGATAGAGACAACATCGTTAGATGTCCTCATCCAGGTACGCAG 808
DB 231847 CGGTACCCCTAGACCAGTGGAAAGCAATATCGTTCGTCTGCTGATGGCGAATTTGGCG 231788

QY 809 AGAAGATGATAGACGCAATCGACAGATACGAGTTTCGAGGGGATTCGGTTCGGTGAAGTGA 868
DB 231787 ACCGCGATGATTGAATTAATCGAAACAAACAGGTAGACAAGGTGATTCTATCGCGGTGTAG 231728
```

```
QY 869 TCACATCGCTCGCTAGAAACGTTCTCTCGGGCTCGGTTCTCTCTGTGTTTCGACAAGCTCG 928
DB 231727 TGGAAATCGTAGCGCGGAATGTTCTCTAAAGGCTTTGGGTGAACCTGTATTCGATAAGTTGG 231668

QY 929 AATCCGAACCTGGCAAAAGCTATGCTTTCTATCTCTCGAGCAACGGGTTTCGAGATTTCGCA 988
DB 231667 AAGCAGATATCGCAAAAGCTGTGATGTCCTTTTACTCTGCTAGCAAGGTTTGAATCGGTT 231608

QY 989 GCGGATTCGCGGAGACCGACTTTCACAGGAAGTGAAGATAATGATAGTATTTATATGATA 1048
DB 231607 CAGGTTTTCGCGGACTCTCTCTAACAGGATTTGAGCATAACGACGAATATATATATTGATG 231548

QY 1049 AGGCTGGAAGTGTACAGACACGGAACCTAATCGCTCGGGTGGTGTGTCAGGGGGGATATCGA 1108
DB 231547 AAAACGGTGAAGTTCGCACACGTAACACCGTTCCGGTGGGATTTCAAGGCGGTATTCGCA 231488

QY 1109 ATGTTGAGATTTGTGCATCTTCAAGTGTCTTTAAAGCGCAGACCATCTATCGGGGTGAAC 1168
DB 231487 ACGGAGAGATATCATTTTCGGAGTTGCATTTAAACCCACAGCACGATTTAGAAAAGAGC 231428

QY 1169 AGAACACCGTGTCAAGGAGCGTCAAGAACGTTGAGCTTCTAGCAAGAGGCGGCATGACC 1228
DB 231427 AAAAAACAGTTACTCTGTGAAGCGAGAAAACCCCTATTAGCCGCCAAAGGTAGACATGATC 231368

QY 1229 CATGCGTCCGCTCGAGCTGTTCTCTGTGGTGAATCCATGGCGCGGTTGTCCTCATGG 1288
DB 231367 CCTGTGTTCTACCCCGTGCAGTCCCAGATGGTAGAAGCAATGGTGGCGGTTGGTACTGTGCG 231308

QY 1289 ACCAGCTGATGGGCGACGTCGGCTCAGTGCAGATGTT 1325
DB 231307 ATCATTTGTTACGCCATCATGGACAGTGAAGGTGCT 231271
```

RESULT 15
AP006487 1/c
WPCOMMENT

Sequence split into 6 fragments LOCUS AP006487 Accession AP006487

Fragment Name	Begin	End
AP006487_0	1	110000
AP006487_1	100001	210000
AP006487_2	200001	310000
AP006487_3	300001	410000
AP006487_4	400001	510000
AP006487_5	500001	523414

Continuation (2 of 6) of AP006487 from base 100001 (AP006487 Cyanidioschyzon merolae str.

Query Match	21.5%;	Score 349;	DB 2;	Length 110000;	
Beat Local Similarity	57.5%;	Pred. No. 1.3e-87;			
Matches	650;	Conservative 0;	Mismatches 475;	Indels 6;	Gaps 1;
QY 211	GCTGCGCTAGAGGTGAAGGCATCTGGAAACACGTTTGGAAACTACTTTTCAGGTTGCAACC	270			
DB 83966	GCTCGTGAGTTGTTTCATGGGCTCGGCACACTCTTTTGGGAGGATATTTTCGAAATACGACT	83907			
QY 271	TATGTTCAATCTCATGGGTTGGTGTGTTGTTTATCATGTTGTTCTCCACCTAGAAAT	330			
DB 83906	TTTGGTGAATCCCAACGTTGGAGTCGGAGTTGTCAATGATGTTGCCACCGGCTG	83847			
QY 331	CCACTCACTGAGGACACCTACAAAGTTGAACTCGATCGAAGACGCGCCGAGACAGACAGA	390			
DB 83846	CCACTTACTATAGAAGACGTTTCAGATTGAGTTAAACAGCGGTAGACCCGGGACAGCTAGG	83787			
QY 391	ATAACCTCCACAAGGAAGGAGACTGTATACATGCAAAATTTCTGTGAGGACACATGAAGGG	450			
DB 83786	CTCAGGACACACAGGAACGAAGCAGATGAAGCGGAGATCCTCTCTGCGCTGTCGACGGT	83727			
QY 451	GTGACTACTGGAACCGCAATTTCTGTTATTGTTCCCAACACAGATCAAAATAGGACAGTAT	510			
DB 83726	GTCACTATGGGTACACCTTATAGCTATCTGTTTTCGAAACACAGACCAAGGCTCTCAGAT	83667			
QY 511	CACGCTGAAATAGCCAAATGTGTACCGCACTTCTCATGACAGACGCAACTTATGACTTCAAG	570			

Db 83666 TATGAGGCAATGAAGTAAGCTATCGACCATCGCAGCGGATGCAACCTATGATGCAAAAG 83607
 Qy 571 TAAGGTGTTAGAGCTGTACAGAGGAGTGGAGGTCTCGGGCAGAAAACCGTTGGAAGG 630
 Db 83606 TACGGCATACGGCGCGTCTCGGTGAGCGTTATCAGCTCGGGAACAATAGGACGT 83547
 Qy 631 GTGGCTGCAAGGGCCCTCCCAAGAAAATTTCTTAAGCTCAAAATGTGGATTTAGAGATCTTG 690
 Db 83546 GTTGTCTGACAGGCCATTGCCAAAAGATCCCTTCGACATATGCGGAACACAAATCTT 83487
 Qy 691 TCGTTTGTTCCTCAAGTGCATCAGGTTGTCTCCAGAAAGACGCGGTTGATTTATGGGTCT 750
 Db 83486 GCCTACGTGAAAAGCGTGAAGCATATCGAG-----GCAGCGAATATTGATCCCAACGTG 83433
 Qy 751 GTAACTTTGGAACAGATAGAGAGCAACATCGTTAGATGTCCTGATCCAGAGTACGCAGAG 810
 Db 83432 GTTTCGTACGAGCAATCGAATCGAACATCGTTTCGCTGCCCTGATTCAGAGGACGATCT 83373
 Qy 811 AAGATGATAGACGCAATCGACAGAGTACGAGTTTCGAGGGGATTCGGTTCGGTGGAGTGATC 870
 Db 83372 CAATGATCCAGCTTATCGATGAATTCGCCGAACAGGTAACAGTGTCCGGCGTGTGTT 83313
 Qy 871 ACATGCTGCTAGAAAAGTTCTCGCGGCTCGGTTCTCTGTGTTCGACAAAGCTCGAA 930
 Db 83312 GAATCGTTTCAAGGAACGTCCTCGCGGCTAGGATCACCGGTGTTTGACAAGCTCGAA 83253
 Qy 931 TCGAACTGGCAAAAGCTATGCTTTCTATTCTCGGAGCAACGCGTTCGAGATTGGCAGC 990
 Db 83252 GCTGAATCGCAAAAGCCCTCATGTCGTTCCAGCATCGAAAGGCTTTGAAGTAGGCTCG 83193
 Qy 991 GGATTCGCGCGGACCGACTTGAACAGAGTGAAGCATATGATGAGTTTATATGATAG 1050
 Db 83192 GGTTCGCGGCGACTTTCGCTAACGGGTGTAGAGCACACGATCCCTTCTATATAGTAGC 83133
 Qy 1051 GCTGGAAGTGTGAGACACGACTAATCGCTCGGTGTGTGTCAGGGAGGATATCGAAT 1110
 Db 83132 AATGTTAGGACCGGAGCGGAGGAGAACCGTTCGGAGGTATTTCAAGGAGGCATCTCGAAC 83073
 Qy 1111 GTTGAGATTGTGCACCTTCAAGCTTCTTTAAGCCGACACCATCTATCGGGGTGAACAG 1170
 Db 83072 GGAGAGAACATTGTATGTGCGCTCGCTTCAAAACCGACCTCTACGATATCAGCGGAACAG 83013
 Qy 1171 AACACCGTGTCAAGGAGCGTCAGAACGTTGAGCTTTCTAGCAAGAGGCGCCATGACCCA 1230
 Db 83012 AATCAGTGAACGCGTCCGGTGAGCAAGTAGTCTTCGCGCAAGGGCAGACACGATCCC 82953
 Qy 1231 TCGGTGCCCCCTCGAGCTGTTCTCTGTGGGAATCCATGGCCGCGTGTGTTCTCATGGAC 1290
 Db 82952 TGTGTGTTGCTCGTGTCTTCCGATGGTAGAAGCCATGGTCGCGCTCGTTCTTGTGTGAT 82893
 Qy 1291 CAdCTGATGGCGCAGCTGGCTCAGTCCGAGATGTTCCGCTCAATACTGCA 1341
 Db 82892 CACTTGTCTCAACAGATTGCCAGGTTCGCTGCTTCCTATTGATCCATCA 82842

Search completed: August 27, 2005, 02:58:23
 Job time : 6886 secs

THIS PAGE BLANK (USPTO)

CC encoding chorismate synthase. This clone is isolated from cs1n cDNA
CC library which is derived from corn silk tissue seedling. Chimeric gene
CC comprising this cDNA operably linked to regulatory sequences is used to
CC transform host cells to alter the level of expression of chorismate
CC synthase. The gene and its products may be used for immunological
CC screening of cDNA expression libraries and to create transgenic plants
CC which may also be herbicide resistant. Synthetic peptides derived from
CC the gene are to raise antibodies, and used in screening assays to
CC identify inhibitors which may be useful as herbicides. Probes and primers
CC from chorismate synthase gene are used for gene mapping and as markers
CC for traits linked to those genes
XX
SQ Sequence 1626 BP; 405 A; 382 C; 439 G; 400 T; 0 U; 0 Other;

Query Match 100.0%; Score 1626; DB 3; Length 1626;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCACGAGCTCAGCTTCGTCCTCTCCGCGCGCGGCAACTATCATCTTCATTAG 60
DB 1 GCACGAGCTCAGCTTCGTCCTCTCCGCGCGCGGCAACTATCATCTTCATTAG 60
QY 61 CTCATCCAACTCTATTCGATGACGACCGTGCACGACGAGCGTGCACG 120
DB 61 CTCATCCAACTCTATTCGATGACGACCGTGCACGACGAGCGTGCACG 120
QY 121 GCACGGCTCGACCCCGCGGATCGGCGCTTGTGGAGTTGCCCHAGCCTCTCTCC 180
DB 121 GCACGGCTCGACCCCGCGGATCGGCGCTTGTGGAGTTGCCCHAGCCTCTCTCC 180
QY 181 CTCGCTTCGCGTGACCGCTGCGCACCTGCTGCGCTAGAGGTGAAGGCAATCGGAAC 240
DB 181 CTCGCTTCGCGTGACCGCTGCGCACCTGCTGCGCTAGAGGTGAAGGCAATCGGAAC 240
QY 241 ACGTTTGAAACTACTTTTCAGGTTGCAACCTATGTTGAATCTCATGGGGTGGTGGT 300
DB 241 ACGTTTGAAACTACTTTTCAGGTTGCAACCTATGTTGAATCTCATGGGGTGGTGGT 300
QY 301 TGTGTTATCAGTGGTTGTCACCTAGAAATTCACCTCAGTGGGAGACCTACAGTTGAA 360
DB 301 TGTGTTATCAGTGGTTGTCACCTAGAAATTCACCTCAGTGGGAGACCTACAGTTGAA 360
QY 361 CTCGATCGAAGAGCGCGCGGACAGAGCAGAAATACCTCCAAAGAGGAGACTGATACA 420
DB 361 CTCGATCGAAGAGCGCGCGGACAGAGCAGAAATACCTCCAAAGAGGAGACTGATACA 420
QY 421 TGCAAAATCTGTTCAGGACACATGAAGGGTGACTACTGAAACGCCAAATCTTTGTTAT 480
DB 421 TGCAAAATCTGTTCAGGACACATGAAGGGTGACTACTGAAACGCCAAATCTTTGTTAT 480
QY 481 GTCCCAACACAGATCAAAATAGGCGAGTGATCACCGTGAATAGCCAAATGTGTACCGACCT 540
DB 481 GTCCCAACACAGATCAAAATAGGCGAGTGATCACCGTGAATAGCCAAATGTGTACCGACCT 540
QY 541 TCTCATGACAGCGCAACTTATGACTTCAAGTACGGTGTAGAGCTGTACAGGAGGTGGG 600
DB 541 TCTCATGACAGCGCAACTTATGACTTCAAGTACGGTGTAGAGCTGTACAGGAGGTGGG 600
QY 601 AGGTCTCCGGCAGAAAAACCGTTGGAAAGGGTGGCTGCAGGGCGCCCTCCCAAGAAAATT 660
DB 601 AGGTCTCCGGCAGAAAAACCGTTGGAAAGGGTGGCTGCAGGGCGCCCTCCCAAGAAAATT 660
QY 661 CTTAAGCTCAAAATGGGATAGAGATCTGTGCTTTGTTTCCAAAGTGCATCAGGTTGTG 720
DB 661 CTTAAGCTCAAAATGGGATAGAGATCTGTGCTTTGTTTCCAAAGTGCATCAGGTTGTG 720
QY 721 CTCCTCAGAGACGCGGTTGATTATGGGTCTGTAACTTTTGGAAACAGATAGAGCAACATC 780
DB 721 CTCCTCAGAGACGCGGTTGATTATGGGTCTGTAACTTTTGGAAACAGATAGAGCAACATC 780
QY 781 GTTAGATCTCCTGATCCAGAGTACGACAGAGAGATGATAGACGCAATCGACAGAGTACGA 840
DB 781 GTTAGATCTCCTGATCCAGAGTACGACAGAGAGATGATAGACGCAATCGACAGAGTACGA 840

QY 841 GTTCGAGGGGATTCGGTTCGGTGGAGTGATCAGATGCGTCCGTAGAAACGTTCTCTCGCGG 900
DB 841 GTTCGAGGGGATTCGGTTCGGTGGAGTGATCAGATGCGTCCGTAGAAACGTTCTCTCGCGG 900
QY 901 CTGGTTCT 960
DB 901 CTGGTTCT 960
QY 961 CTGCGAGCAACCGGTTTCGAGATTGGCAGCGGATTCGCCGGGACCGACTTTCACAGGAAGT 1020
DB 961 CTGCGAGCAACCGGTTTCGAGATTGGCAGCGGATTCGCCGGGACCGACTTTCACAGGAAGT 1020
QY 1021 GAGCATATGATGAGTTTATATGGAATGAGGCTGGAAGTGTCCAGGACACCGACTTAATCGC 1080
DB 1021 GAGCATATGATGAGTTTATATGGAATGAGGCTGGAAGTGTCCAGGACACCGACTTAATCGC 1080
QY 1081 TCGGGTGCTGTGACGAGGAGGATATCGAATGTTTTCAGATTGTGCACATTCCTCAAGTTTCTTTT 1140
DB 1081 TCGGGTGCTGTGACGAGGAGGATATCGAATGTTTTCAGATTGTGCACATTCCTCAAGTTTCTTTT 1140
QY 1141 AAGCCGACACCATCTATTCGGGGTGAACAGAAACACCGTGTCAAGGGAGCGTCAAGACGTT 1200
DB 1141 AAGCCGACACCATCTATTCGGGGTGAACAGAAACACCGTGTCAAGGGAGCGTCAAGACGTT 1200
QY 1201 GAGCTTCTAGCAAGAGGCGCCCATGACCCATGCGTCCGCTCGAGCTGTTCTCTGTGGTG 1260
DB 1201 GAGCTTCTAGCAAGAGGCGCCCATGACCCATGCGTCCGCTCGAGCTGTTCTCTGTGGTG 1260
QY 1261 GAATCCATGCGCGTTCCTCATGGAACAGCTGATGGCGACCGTGGCTCAGTGGCGAG 1320
DB 1261 GAATCCATGCGCGTTCCTCATGGAACAGCTGATGGCGACCGTGGCTCAGTGGCGAG 1320
QY 1321 ATGTTTCGCTCAATCTGACCTTCAAGAACCGAGTTGGCTCTTTCTAGCAGAGCGAGC 1380
DB 1321 ATGTTTCGCTCAATCTGACCTTCAAGAACCGAGTTGGCTCTTTCTAGCAGAGCGAGC 1380
QY 1381 ACACCTGATGAGCTCGCGCCAAATTTATCATTTATCATAGTAATAGTAGCTCAAGCGT 1440
DB 1381 ACACCTGATGAGCTCGCGCCAAATTTATCATTTATCATAGTAATAGTAGCTCAAGCGT 1440
QY 1441 GGCTTGGTTTGTCTCTTTCGACCGTAGTTTTTTTTTTTTTTTTTTTTTCCCGAAGTGTGATG 1500
DB 1441 GGCTTGGTTTGTCTCTTTCGACCGTAGTTTTTTTTTTTTTTTTTTTTTCCCGAAGTGTGATG 1500
QY 1501 CGATGAAGTGAATAAGGCACTTGGTTTCCTGTGTCATTTGTACACGTTTTCATATAATGTAA 1560
DB 1501 CGATGAAGTGAATAAGGCACTTGGTTTCTGTGTCATTTGTACACGTTTTCATATAATGTAA 1560
QY 1561 TCTACTTTCGAGATGATGCAATTTTATAGATGTTGGCTTGTGAAGACAAAAA 1620
DB 1561 TCTACTTTCGAGATGATGCAATTTTATAGATGTTGGCTTGTGAAGACAAAAA 1620
QY 1621 AAAAAA 1626
DB 1621 AAAAAA 1626

RESULT 2
AAZ50435
ID AAZ50435 standard; cDNA; 1635 BP.
XX
AC AAZ50435;
XX
DT 18-MAY-2000 (first entry)
XX
DE Corn clone Chpc24.pk0002.h1.fis encoding chorismate synthase.
XX
KW Corn cDNA clone chpc24.pk0002.h1.fis; chorismate synthase;
KW immunological screening; herbicide resistance; antibody; gene mapping;
XX corn; ss.
OS Zea mays.

XX Key Location/Qualifiers
 FH 52..1374
 FT /*tag= a
 FT /product= "Chorismate synthase"
 XX WO200005353-A2.
 XX 03-FEB-2000..
 XX 20-JUL-1999; 99WO-US016353.
 XX 21-JUL-1998; 98US-0093611P.
 XX (DUPO) DU PONT DE NEWMOURS & CO E I.
 XX Cahoon RE, Falco SC, Pember SO;
 XX WPI; 2000-182687/16.
 XX DR P-PSDB; AAY44890.
 XX
 XX New chorismate synthase polypeptides used to alter the level of the
 XX enzyme and thus the level of aromatic to non aromatic amino acids in
 XX transformed plants.
 XX
 XX Claim 3; Page 28; 39pp; English.
 XX
 XX The present sequence is corn cDNA clone designated chpc24.pk0002.h1.fis
 XX encoding chorismate synthase. This clone is isolated from chpc24 cDNA
 XX library which is derived from corn 8 day old shoot treated 24 hours with
 XX herbicide. Chimeric gene comprising this cDNA operably linked to
 XX regulatory sequences is used to transform host cells to alter the level
 XX of expression of chorismate synthase. The gene and its products may be
 XX used for immunological screening of cDNA expression libraries and to
 XX create transgenic plants which may also be herbicide resistant. Synthetic
 XX peptides derived from the gene are to raise antibodies, and used in
 XX screening assays to identify inhibitors which may be useful as
 XX herbicides. Probes and primers from chorismate synthase gene are used for
 XX gene mapping and as markers for traits linked to those genes
 XX
 XX Sequence 1635 BP; 441 A; 367 C; 422 G; 405 T; 0 U; 0 Other;
 XX
 XX Query Match 53.5%; Score 870.4; DB 3; Length 1635;
 XX Best Local Similarity 80.5%; Pred. No. 1.6e-240;
 XX Matches 1031; Conservative 0; Mismatches 246; Indels 3; Gaps 1;
 XX
 XX 83 CGACCGTCCCAAGCCACAGCAGGTGGCGCATCTACGGGCGCGTCCACACGGTTCTCCCGCGCGGA 142
 XX 59 CGCCCGTGTCCGAGCGCGCGGTGTCCGCGGCGGTCCACACGGTTCTCCCGCGCGGA 118
 XX 143 TCGCGCCCTGTGGAGTTTGGCCCGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 202
 XX 119 TACGGCGCTTCCCGAGTCCGCCC---CCAGCTCCCTCCGGTTATCCGTCGGCCCGCGTC 175
 XX 203 GCGCACCTGTCTCGCTAGAGGTGAAGGCATCTTGAAACACACGTTTGGAACTACTTTTCAGG 262
 XX 176 GCGCGCGCTCCAGCTAGAGGTGAAGGCATCAGGAAATGTGTTCCGGAATCTTCCAGG 235
 XX 263 TTGCAACCTATGGTGAATCTCATGGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 322
 XX 236 TTGCAACCTATGGTGAATCTCATGGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 295
 XX 323 CTGGAATCTCACTAGGCGAGACCTACAGTTGAACTCGATCGAGACGGCCCGGAC 382
 XX 296 CAGAAATCTCTCTACTAGGCGAGACATGCAAGTAGAATCGATAGAGACGTCCCGGTC 355
 XX 383 AGAGCAGAAATACCTCCCAAGGAGGAGACTGATACATGCAAAATCTGTCCAGGACAC 442
 XX 356 AAGTAGAATACACCCCAAGAGGAGACTGATACATGCAAAATCTGTCCAGGACAC 415
 XX 443 ATGAAGGGGTGACTACTGGAACCGCAATTTCTTTTATTTGTCCTCAACACAGATCAATAG 502
 XX 416 ATGATGGGATGACTACTGTCACCAATTCAGTCTTTGTTGCCAACAACACAGATCAAGGG 475

QY 503 GCAGTGATCAACCGTGAATAGCCATGTGTACCGACCTTCTCATCGACGCAACTTATG 562
 DB 476 GTGGTGATACAGTGAATGTCTAAGGCGTACAGACCATCCCATGCAGATGCAACTATG 535
 QY 563 ACTTCAAGTACCGTGTAGAGCTGTACAGGAGGTGGAGGTCTCTCGGCGACAAAACCG 622
 DB 536 ACTTCAAGTATGAGCTGTAGAGCTGTGACAGGAGGTGGAAGGTTCATCAGCAGAGAACCA 595
 QY 623 TTGGAAGGTGGCTGCAGGGCCCTCCCAAGAAATCTTAAAGCTCAAAATGTGATAG 682
 DB 596 TTGGCAGGGTGGCTGCAGGAGCTCTTGTCAAAAGAAATCTTAAAGCTCAAAATCAGGAGTGG 655
 QY 683 AGATCTTGTCTGTCTTGTCTTCAAGGTGCATCAGGTGTGTCTCCCAAGACGCGGTGTGATT 742
 DB 656 AGATCTTGTCTTGTCTTCAAGGTGCATCAGGTGTGTCTCCCAAGACGAGTGTGATT 715
 QY 743 ATGGGTCTGTAACTTTTGGAAACAGATAGAGACCAATCGTTAGATGTCTCTGATCCAGAGT 802
 DB 716 ATGAGACTGTAACTTTTGGAAACATATAGAGACCAATCGTTAGATGTCTCTGATCCAGAT 775
 QY 803 ACGCAGAGAAATGATAGACCAATCGACAGATACGAGTTCGAGGGGATTCGCTCGGTG 862
 DB 776 ATGCAGAGAGATGATTCCTGCCATTGATACGGTACGAGTTAGAGAGATTCATTTGGTG 835
 QY 863 GAGTGATCACATCGCTCGCTAGAACGTTCTCTCGGGCTCGGTCTCTCTGTGTTGACA 922
 DB 836 GGGTGTGATGATGATTCGAAGAAATGTCTCTCGGTGCTCTCTGTGTTGACA 895
 QY 923 AGCTCGAACTCGAACTGCGCAAAAGCTATGCTTTTCTTATTCCTCGGAGCAACGGGTTCGAGA 982
 DB 896 AACTTGAAGCTGAACTGCGCAAAAGCTATGCTTTCTTCTTCTGCAAGCAAGGGTTTGAGA 955
 QY 983 TTGGCAGCGGATTCGCGGGGACCGACTTTGACAGGAAGTGAGCATTAATGATGATTTTATA 1042
 DB 956 TTGGCAGTGGGTTCGCTGCTGCGGACTTTTACTGGAAGTGAGCATAATGATGATTTCTATA 1015
 QY 1043 TGGATAGCTGGAAGTGTGAGGACGAGTAACTCGCTCGGTGGTGTGTCAGGAGGGA 1102
 DB 1016 TGGATGAGGCTGGAAGTGTGAGGACACGAATTAATCGCTCAGGGCGGTGTTGAGGAGGGA 1075
 QY 1103 TATCGAATGTTGAGATGTGCACTTTCAAGTTGCTTTTAAAGCCGACACCATCTATCGGGG 1162
 DB 1076 TATCAATGTTGAAATTAATTTACTTTCAAGTGGCTTTTAAAGCCACCACTATCGGA 1135
 QY 1163 TGAACAGAAACACCGTGTCAAGGAGCGTCAAGACGTTTGAGCTTCTAGCAAGAGGCGCC 1222
 DB 1136 AGAAGCAAAATACTGTGTCAAGGAGGACATGAGGATGTTGAACCTTTTGGCAAGGGGCGCC 1195
 QY 1223 ATGACCCATGCTCGCCCTCGAGCTGTTCTGTGGTGGAAATCCATGGCGGTGGTCC 1282
 DB 1196 ATGACCCCTGTGTTGTTCTCTCGAGCTGTTCTTATGTTGGAAATCCATGGCTGGTCC 1255
 QY 1283 TCATGGACCACTGATGGCGCACGCTGCTCAGTGCAGATGTTGCGCTCAATATCTGCAC 1342
 DB 1256 TGATGGACCACTCATGGCGCATATGCGCAGTGTGAGATGTTTCCGCTGAACCTTGCC 1315
 QY 1343 TTCAAGAACCAAGTTGGGTCT 1362
 DB 1316 TACAAGAGCCCATTTGGGTCT 1335

RESULT 3
 AAZ50440

ID AAZ50440 standard; cDNA; 966 BP.

XX AC AAZ50440;

XX 18-MAY-2000 (first entry)

DE Rice clone rls72.pk0029.g8 encoding chorismate synthase.

KW Rice cDNA clone rls72.pk0029.g8; chorismate synthase;

immunological screening; herbicide resistance; antibody; gene mapping;
rice; ss.

Oryza sativa.

Key Location/Qualifiers
FH 45..964
CDS /*tag= a
FT /product= "Chorismate synthase"
FT /note= "no stop codon given"

WO200005353-A2.

03-FEB-2000.

20-JUL-1999; 99WO-US016353.

21-JUL-1998; 98US-0093611P.

(DUPO) DU PONT DE NEMOURS & CO E I.

Cahoon RE, Falco SC, Pember SO;
WPI; 2000-182687/16.
P-PSDB; AAY44895.

New chorismate synthase polypeptides used to alter the level of the
enzyme and thus the level of aromatic to non aromatic amino acids in
transformed plants.

Claim 3; Page 35; 39pp; English.

The present sequence is rice cDNA clone designated r1e72.pk0029.98
encoding chorismate synthase. This rice is isolated from r1e72 cDNA
library which is derived from rice leaf 15 days after germination, 72
hours after infection of strain Magaporthe grisea 4360-R-67. Chimeric
gene comprising this cDNA operably linked to regulatory sequences is used
to transform host cells to alter the level of expression of chorismate
synthase. The gene and its products may be used for immunological
screening of cDNA expression libraries and to create transgenic plants
which may also be herbicide resistant. Synthetic peptides derived from
the gene are to raise antibodies, and used in screening assays to
identify inhibitors which may be useful as herbicides. Probes and primers
from chorismate synthase gene are used for gene mapping and as markers
for traits linked to those genes

Sequence 966 BP; 251 A; 234 C; 264 G; 217 T; 0 U; 0 Other;

Query Match 38.2%; Score 621; DB 3; Length 966;
Best Local Similarity 80.4%; Pred No. 1.6e-168;
Matches 742; Conservative 0; Mismatches 175; Indels 6; Gaps 1;

75 TCCGATGACGACCGTGCACCAAGCCACAGCAGGTGCGCACTCAGCGGCACGGCTCGCAC 134
41 TCCGATGCGCGCCACCAAGTGTGTCAGCGGTGCGCGCGTCTCCCTCCCGCGCGGG 100
135 CGCGCGATCGCGCGCTTCTGGAGTTTGGCCAGCCTCTCCCTCCCTCCCTCGCGGT 194
101 CGCGGGGTTCGCGCGCTTCCCGGAGTTCGCGCGCGCT-----TCCCTCCGCTTCTCGGT 154
195 GCACCGCTGCGCGCACTGTCTGCTAGAGGTGAAGGCATCTGMAACAGTTTGGAAACTA 254
155 CGGCGCGCGCGCGCGCTGCGCTAGAGGTGAAGGCGTCTGCAATGTTTGGAACTA 214
255 CTTTCAGGTTGCAACTATGTTGAATCTCATGGGGGTGGTGTGGTGTGTTATCAGTGG 314
215 CTTTCAGGTTGCAACTATGTTGAAGTCTCATGGAGGCTCATGGAGCGGTGGTTCGCTAATCAGTGG 274
315 TTGTCACCTAGTAATTCCTCACTCAGCGCAGACCTACAAGTTGAACTCGATCGAAGCG 374
275 ATGCCACCCAGAAATCCCACTTACTGAAAGCAGATATGCAAGTAGAACTCGACCGGAGCG 334
375 GCCCGGACAGACGAGAATAAACCCTCCCAAGGAAGGAGACTGATACATGCAAAATTTCTGTC 434

Db 335 GCCAGGCAGCAGCAGATACCAACCCCAAGAGGAGACTGACACTTGCAAAATTTCTTC 394
Qy 435 AGGACACATGAAGGGGTGACTACTGGAAAGCCAAATCTTGTATTGTGTCCCAACACAGA 494
Db 395 AGGACACATGAAGGAATGACCACTGGGACACCAATTCATGTGTTTGTCCGGAACAGA 454
Qy 495 TCAATAGGCAGTGATCACCGTGAATAGCAATGTGTACCGACCTTCTCATCAGACGC 554
Db 455 TCAGAGAGGGGTGATTACAGTGAATGGCTAAGGCCCTACAGACCTTTCATCATGAGATGC 514
Qy 555 AACTTATGACTTCAAGTACCGTGTAGAGCTGTACAGGAGGTGGAGGCTCTCGGCGAG 614
Db 515 AACTTATGACTTCAATACCGTGTAGAGCAGTGCAGGAGGTGGAAGATCATCAGCAAG 574
Qy 615 AAAAACCGTTGGAAGGGTGTGTCAGGGGCGCTCCCAAGAAATTTCTTAAGCTCAAAATG 674
Db 575 AGAGACCATTTGGAAGGGTGTGTCAGGAGCTCTTGCAGAAAGAAATTTCTTAAGCTCAAAATC 634
Qy 675 TGGATTAGAGATCTTGTGCTTTGTTTCCAAAGTGCATCAGGTTGTGCTCCCAAGACGC 734
Db 635 TGGAGTAGAGATCTTGGCGTTTGTGCCAAGGTGCATCAAGTTGTACTACAGAAAGATGC 694
Qy 735 GGTTGATTATGGGTCTCTAACTTTGGAACAGATAGAGCAACATCGTTAGATGCTCTGA 794
Db 695 CGTTGATTATGACACTGTAACTGGAACAGATAGAAAGCAACATTTGTAGATGCTCTGA 754
Qy 795 TCAGAGTAGCAGAGAGATGATAGACGCAATGCACAGAGTAGCAGTTGAGGGGATTC 854
Db 755 TCCAGAAATATGCACAGAAAGATGATTGATGCACTCGATAAAAGTACGAGTTAGAGTGATTC 814
Qy 855 GGTGCGTGGAGTGCATGCATGCGTCCGTAGAAAGTTCTTCGCGGGCTCGGTTCTCCTGT 914
Db 815 GATTGGTGGTGTGTCATGTCATTCGCAAGAAATGTTCTTCGCGGATTTGGCTCTCCTGT 874
Qy 915 GTTCGACAAAGCTCGAATCCGAACTGGCAAAAGCTATGCTTTCTATTCTCGCAGCAACGG 974
Db 875 ATTTGACAAACTTGAGGCTGAATTTGGCGAAAGCTATGCTTTCTCTCTCTCAAGCAAGGG 934
Qy 975 GTTCGAGATTGGCAGCGGATTCG 997
Db 935 GTTTGAGATCGCAGTGGATTG 957

RESULT 4
AAC40227
ID AAC40227 standard; DNA; 1595 BP.
XX AC AAC40227;
XX AC AAC40227;
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 27533.
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
PF 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 28-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 24-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142300P.
PR 09-JUL-1999; 99US-0142803P.
PR 12-JUL-1999; 99US-0142920P.
PR 13-JUL-1999; 99US-0142977P.
PR 14-JUL-1999; 99US-0143545P.
PR 15-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 26-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158023P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.

XX This invention relates to a novel method for altering one or more plant
CC characteristics. Specifically, it refers to identifying genes that are up
CC - or down-regulated in transgenic plants overexpressing the heterodimeric
CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
CC alter plant characteristics accordingly. The present invention describes
CC generating transgenic plants for the production of growth regulators,
CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
CC the altered plant characteristics are selected from increased yield or
CC biomass, enhanced survival capacity, stress tolerance, plant architecture
CC or physiology, altered endoreduplication, biochemistry, signal
CC transduction, storage lipid mobilization and/or altered photosynthesis,
CC each relative to the corresponding wild type plants. Accordingly, these
CC sequences can also be useful as positive or negative selectable markers
CC during transformation of cells or tissues. The identified genes play a
CC role in a variety of biological processes such as DNA replication, cell
CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as
CC transcription factors. This polynucleotide sequence is the cDNA
CC transcribed from a 1.3 fold or more in plants overexpressing the E2Fa/Dpa
CC transcription factor, given in an exemplification of the invention.
XX
SQ Sequence 1311 BP; 369 A; 277 C; 318 G; 347 T; 0 U; 0 Other;

Query Match 32.9%; Score 535.4; DB 12; Length 1311;

Best Local Similarity 66.9%; Pred. No. 1.1e-143;

Matches 761; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

QY 216 CCTAGAGGTGAAGGCATCTGGAACACAGTTCGAACTACTTTCAGGTGGCAACCTATGG 275
DB 135 CTTCCAGATACAAAGTACTGGAAGTTTCATATGGGACTATTTTCGAGTTTCAACTTTGG 194
QY 276 TGAATCTCATGGGGTGGTGTGGTGTATTCAGTGTGTCCACCTAGATATCCACT 335
DB 195 AGATYCATATGAGGAGGAGTGGTGTATCATATGATGTTGTCTCTCGTATTCCTACT 254
QY 336 CACTGAGGAGCCTCAAAAGTTGAATCGATCGAAGACGGCCCGGACAGACAGAAATAC 395
DB 255 TACTGAATCTGATTTGCAATTCGATCTCGATAGAGGGGCTGGTCAGACGAGGATCAC 314
QY 396 CTTCCACAGGAGGAGATGATACATGCAAAATTCGTTCAGGACACATGAAGGGTGAC 455
DB 315 AACTCCTAGAAAGAGACTGATCTTGGCGGATATCGTCTGGAGTCTCTGAAGGAATGAC 374
QY 456 TACTGGAACGCCAATCTTGTGTATTTGTCTCCAAACACAGATCAAAATAGGACGATCACCG 515
DB 375 GACAGGACACCTATCCATGTGTGTATCCAAACACAGATCAGAGGACTTGTATACG 434
QY 516 TGAATATAGCAATGTGTACCGACCTTCTCATGACGACGCAACTTATGACTTCAAGTACGG 575
DB 435 TGAATATGTCGGTTGCTATAGACCATCGCATGCTGATGCACTTATGACATGAATGATGG 494
QY 576 TGTATAGAGCTGTACAGGAGGTGGGAGTCTCGGACAGAAAACCGTTGGAAGGGTGGC 635
DB 495 TGTGATAGTGTGACAGGGTGGGAGAGATCTTCAGCTAGAGAGACCATTTGGAAGAGTTGC 554
QY 636 TGAGGGGGCCCTCCCAAGAAAATCTTAAGTCTCAAACTGTGATAGAGATCTGTCTGTT 695
DB 555 TCTGTGAGCTTTGGCCAGAAAATTTTGAAGCAATTTGAGGAACCTGAGATTTCTGCTTA 614
QY 696 TGTTTCCAAAGTGCATCAGGTGTGTCTCCAGAAAGCGGGTTGATTTATGGTCTGTAAAC 755
DB 615 TGTCTCGAAGTTTCCATGTTGTACTTCCAGAAAGATTTGGTAGACACAGAAATTTAAC 674
QY 756 TTTGGAACAGATAGAGAGCAATCGTTAGATGTCTGATCAGAGTACGACAGAGAT 815
DB 675 ACTCGAACAGATAGAAAATAACATTTGTTCAGATGCCCTTAATCCGAGTATGCGGAAAGAT 734
QY 816 GATAGACGAATCGACAGATGAGTTCGAGGGATTCGGTGGTGGATGATCACATG 875
DB 735 GATAGCTGATGATGTGTGTCAGGACAAAAGGAACTCTGTGTGGTGGTGTGACCTG 794
QY 876 CGTCTGTAGAAAGCTTCTCGCGGGCTCGGTCTCTGTGTTCGACAAAGCTCGAATCCGA 935

DB 795 CATTTGTCGGAATGCTCCACGTGGGCTTGGTACACGGTTTTTCGATAAACTTTGAAGCAGA 854
QY 936 ACTGGCAAAAGCTATGCTTTCTATTCTCGCGAGCAACGGTTTCGAGATTGGCAGCGGATT 995
DB 855 ACTGGCAAAAGCTTGTATGTGCTACCTGCAACTAAGGGATTTGAGTTTGGAAAGCGGCTT 914
QY 996 CGCCGGGACCGACTTTCAGCAGGAAGTGAGCATATGATGAGTTTATATGGAATAAGGCTGG 1055
DB 915 TCCAGGTACTTTTTCGACTGGTCTTGAACACATGATGATGTTCTATACCGATGAANAATGG 974
QY 1056 AAGTGTCCAGGACCGGACTTAATTCGTCTGGGTGGTGTGAGGAGGGATATCGAAATGTGA 1115
DB 975 AAGAATATGCTTACCAAGCAACCGGATCTGGTGAATTCAGGAGGGATCTCAAAATGGTGA 1034
QY 1116 GATTGTGCACATTCAAAGTTGCTTTTAAGCCGACACATCTATCGGGGTGAAACAGACAC 1175
DB 1035 AATAATAACATGAGATGAGCTTCAAGCCAAACATCAACAATTTGGAAGGAAGCAAAATAC 1094
QY 1176 CGTGTCAAGGGGAGCGCTCAGAACGTTTGAGCTTCTAGCAAGAGGGCGCCCATGACCCATGCC 1235
DB 1095 GGTAAACAGACAGACAGGTAGAAACCGAATGATTTGGCGGTGGTCTCATGATCTCTGTGT 1154
QY 1236 CGCCCTCTCGAGCTGTCTCTGTGGTGAATTCATGCGCGGTTGGTCTCTCATGAGCAGCT 1295
DB 1155 TGTTCACGAGCTGTGCCAATGGTGGAAAGCAATGGTGGCTCTAGTTCTTGTGGATCAATT 1214
QY 1296 GATGCGGCACGTGGCTCAGTGGAGATGTTTCGGCTCAATCTGACTGCACTTCAAGAAC 1352
DB 1215 GATGCGGCAATACGCAACAATGCCATTTGTTCCAAATNAATCCAGAGTTGCGAGGAACC 1271

RESULT 6

ADT91426

ID ADT91426 standard; DNA; 1311 BP.

XX AC ADT91426;

XX DT 16-DEC-2004 (first entry)

XX DE Arabidopsis thaliana gene 1456.

XX KW Herbicidal; plant growth; plant development; mouse-ear cross; plant;

XX KW gene; ds.

XX OS Arabidopsis thaliana.

XX FH Key Location/Qualifiers

XX CDS 1..1311

XX FT /*tag=

XX FT /product= "protein essential for plant growth and

XX FT development"

XX PN US2004191850-A1.

XX PD 30-SEP-2004.

XX PF 29-OCT-2003; 2003US-00696616.

XX PR 04-NOV-2002; 2002US-0423519P.

XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX PI Levin JZ, Tossberg J, Zhou Q, Aux G, Mcelver J, Thomas C;

XX PI Patton DA, Frye C, Budziszewski G;

XX DR WPI; 2004-698717/68.

XX DR P-PSDB; ADT91427.

XX DR Identifying an herbicidal compound comprises combining a protein
XX essential for growth of Arabidopsis thaliana, with a compound to be
XX tested for the ability to bind to, or inhibit the activity of the
XX polypeptide.

Example 9; SEQ ID NO 27; 131pp; English.

The invention relates to a method for identifying herbicidal compounds, which involves combining a protein essential for growth of *Arabidopsis thaliana* with a compound to be tested for the ability to bind to or inhibit the activity of the polypeptide. The invention also relates to nucleic acid sequences encoding proteins essential for plant growth and development. The methods, nucleic acid molecules and proteins of the invention are useful for identifying herbicidal compounds and for killing or inhibiting the growth or viability of the plant. The present sequence is the *Arabidopsis thaliana* gene encoding protein essential for plant growth and development.

SQ Sequence 1311 BP; 369 A; 277 C; 318 G; 347 T; 0 U; 0 Other;

Query Match 32.9%; Score 535.4; DB 13; Length 1311;
Best Local Similarity 66.9%; Pred. No. 1.1e-143;

Qy	216	CCTAGAGGTGAAGGATCTGAAACA	CGTTTGGAAA	CTACTTT	CAGGTTGCAACTATGG	275
Db	135	CTTCCAGATACAAGCTACTGGAAGTT	CATATGGGACTCA	TTTTCGAGTTTCAACTTTTGG	194	
Qy	276	TGAATCTCATCGGGGTGGTGGTTGTGTTAT	CAGTGGTTGTC	CCACTAGAAATTC	335	
Db	195	AGAAATCACATGGAGGAGGTGGTTGTATCA	TTTGTAT	CTTCTCGTATTC	254	
Qy	336	CACGTGAGGCAGACCTTACAAGTTGAAC	TGATCGAAGACGGCCCGGACAGAGCAGAA	TAAAC	395	
Db	255	TACTGAAATCTGATTTGCAATTCGATCT	CGATAGAGGAGGCGCTGGT	CAGAGCAGATCAC	314	
Qy	396	CTCCACAAGGAAGAGAGACTGATACATG	CAAAATTTCTGT	CAGGACACATGAAGGGGTGAC	455	
Db	315	AACTCTAGAAAGAAGACTGATCTTGCCG	GATATCGTCTG	GAAGTCTCTGAAGGAA	374	
Qy	456	TACTGGRAACGCAATTTCTTGTATTGTG	TCCCAAACACAGATCAAA	TATAGGCAGTGTACCG	515	
Db	375	GACAGGAACACCTATCCATGTGTTGTAT	CCAAACACAGATCAGAGAGGACTT	TGATTACAG	434	
Qy	516	TGAATATAGCCAAATGTGTACCGACCTT	CTCATGCAGAGCAACTTAT	GACTTCAAGTACGG	575	
Db	435	TGAATGTGCGTTGCCATATAGACCAT	CGCATGCTGATGCAACTTAT	GCATGAATATGG	494	
Qy	576	TGTTAGAGCTGTACAGGAGGTGGGAGT	CTCCGGCAGAAAAACCGTTG	GAAGGGTGGC	635	
Db	495	TGTACAGTCACTGCAGGGTGGAGAGAT	CTTCAAGCTAGAGAGACCATT	TGGAAGAGTTGC	554	
Qy	636	TGCGAGGGCCCTCCCAAGAAAAATTTCT	TAAAGTGTGGATTAGAGATCTT	TGTGCTGT	695	
Db	555	TCCTGGAGCTTTGGCCAAAGAAAAATTT	TGAAGCAATTTG	CAGGAACTTGATTTTGCCTA	614	
Qy	696	TGTTTCCAAAGTGATCAGGTGTGCTCC	CAAGAGACCGGTTCTAT	TGGTCTGTAAAC	755	
Db	615	TGTTCTCGCAAGTTCAACATGTTGAT	CTTCCAGAGAAATTTGGTAGACCA	CGAGAAATTTAAAC	674	
Qy	756	TTTGGAAACAGATAGAGACCAACATCG	TATAGTGTCTTGATCCAGAGTACG	CAGAGAAAGAT	815	
Db	675	ACTCGAACAGATAGAAATAACATGTG	CAGATGCCCTAATCCCGAGTAT	TGCGGAAAGAT	734	
Qy	816	GATAGACGCAATTCGACAGAGTACGAG	TTTCGGTTCGGTGGAGTGATCA	CATG	875	
Db	735	GATAGTGCAGATTGATGCTGTGAGCA	CAAAAGGGAACTCTCTTGGTGGT	TGTGTGACCTG	794	
Qy	876	CGTGCCTAGAAACGTTCTCCGGGGCT	CGGTTCTCTGTGTTGCACAAGCT	CGAATCCGA	935	
Db	795	CAITGTTTCGGAATGCTCCACGTGGCT	TGGTACACCGGTTTTCGATAAAC	CTTTGAAGCAGA	854	
Qy	936	ACTGGCAAAAGCTATGCTTTTCTATT	TTCTCTGCGAGCAACGGGTTTC	GAGATTTGGCAGCGGATT	995	
Db	855	ACTGGCAAAAGCTTGTATGTCGCTAC	CTGCAACTTAAGGGAATTTGAGTTT	TGGAACGCGCTT	914	
Qy	996	CGCCGGGACCGACTTGACAGGAAGT	AGACATAATGATGAGTTTATAT	TGGATAAGGCTGG	1055	

CC peptides derived from the gene are to raise antibodies, and used in
 CC screening assays to identify inhibitors which may be useful as
 CC herbicides. Probes and primers from chorismate synthase gene are used for
 CC gene mapping and as markers for traits linked to those genes
 XX
 SQ Sequence 1015 BP; 283 A; 213 C; 251 G; 268 T; 0 U; 0 Other;

Query Match 30.8%; Score 500.6; DB 3; Length 1015;
 Best Local Similarity 80.3%; Pred. No. 1.1e-133;
 Matches 587; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 631 GTGCTGAGGGGGCTCCCAAGAAATCTTAAGCTCAAAATGCTAGAGATCTTG 690
 DB |||||
 QY 5 GAGGCTGAGGAGCTGTTCAGAAATCTTAAGCTCAAAATGCTAGAGATCTTA 64
 DB |||||
 QY 691 TCGTTTGTTCCTCAAGTGCATCAGTGTGCTCCCAAGACGCGGTGATTATGGTCT 750
 DB |||||
 QY 65 GCATTTGTTTCCAAAGTGCATCAAGTGTGCTCCCAAGACGCGGTGATTATGAACT 124
 DB |||||
 QY 751 GTAACTTTGGAACAGATAGAGCAACATCGTTAGATGCTCTGATCCAGAGTACGAG 810
 DB |||||
 QY 125 CTTACCTGGATCAGATAGAGCAACATTTGTAGATGCTCTGATCCAGATATGACAG 184
 DB |||||
 QY 811 AAGATGATACCGAATCGACAGATACGAGTTCGAGGGATTCGTCGTCGAGTGATC 870
 DB |||||
 QY 185 AAGATGATGATGCAATTCGATAAAGTACGAGTTCGAGGGATTCGATTCGTCGTCG 244
 DB |||||
 QY 871 ACATGCTGCTAGAAAGCTTCCTCGGGCTCGTTCCTGCTGATTCGACAGTCTGAA 930
 DB |||||
 QY 245 ACATGCAATTCGCAAGAAATGTTCTCGTGGCTTGGCTCTCTCTGATTTTACAAACTGAA 304
 DB |||||
 QY 931 TCGAATGTCGCAAAAGCTATGCTTTCTTATCTCTCGAGCAACGGGTTCGAGATTCGACG 990
 DB |||||
 QY 305 GCTCTACTGCAAGGCTATGCTTTCTCTCTGCAAGCAAGGGTTCGAGATTCGTAAT 364
 DB |||||
 QY 991 GGATTCGCGGACCGACTTGACAGCAAGTGCATTAATGATGAGTTTATATGATAAG 1050
 DB |||||
 QY 365 GGATTTGCAAGTACTGACCTAACTGGAAGTGCATTAACGATGAGTTCTATATGACGAG 424
 DB |||||
 QY 1051 GCTGGAAGTGTGAGGAGCGGCTATCTCTCGGGTGTGTCGAGGAGGATTCGAT 1110
 DB |||||
 QY 425 GCTGGAATGTAAGACAGCAACCAATCGCTCGGGGCTGTACAGGAGGATTCAAAT 484
 DB |||||
 QY 1111 GTTGAATGTGACTTCAAGTGTCTTTTAAGCCGACACCACTATCTATCGGGTGAACAG 1170
 DB |||||
 QY 485 GGTGAATATATATCTTCAAGTAGCTTTCAAGCCCAAGCACTATTGGGAAGCA 544
 DB |||||
 QY 1171 AATACGCTGTCAAGGAGCGTCAGAACGTTGAGCTTCTAGCAAGAGGCGCCATGACCA 1230
 DB |||||
 QY 545 AATCTGTAAACAGGATCATGAGGATTCGAACTTCTGCAAGGGGTTCGCCATGACCA 604
 DB |||||
 QY 1231 TGGTTCGCCCTCGAGCTGTTCTGTTGGTGAATCCATGGCGGCTTGGTCTCATGAC 1290
 DB |||||
 QY 605 TGTGTCGTCCTCGGCTGTTCCATGTTGAGAGATGCTGATGCTTCTCATGAC 664
 DB |||||
 QY 1291 CACTGATGCGGACGTCGCTGATGCGGATGTTTCGCTCAATATCTGCACTTCAAGAA 1350
 DB |||||
 QY 665 CAGCTGATGGCACATGTTGCTCAGTGCAGATGTTCCGCTGAACTTCGCTTCAAGAA 724
 DB |||||
 QY 1351 CCAGTTGGCTC 1361
 DB |||||
 QY 725 CCATCGGCTC 735

RESULT 8
 ABX48414
 ID ABX48414 standard; cDNA; 432 BP.
 XX
 XX AC ABX48414;
 XX
 XX
 XX
 XX 21-FEB-2003 (first entry)
 XX
 DE Bovine EST associated with lactation/muscle/fat deposition #13579.

XX
 KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
 KW muscle deposition; fat deposition; genome mapping; gene identification;
 KW gene analysis; cattle breeding.
 XX
 OS Bos Taurus.
 XX US2002137139-A1.
 XX 26-SEP-2002.
 XX 24-SEP-2001; 2001US-00960352.
 XX 12-JAN-1999; 98US-0115707P.
 XX 11-JAN-2000; 2000US-00480902.
 XX (BYAT/) BYATT J C.
 XX (MATH/) MATHIALAGAN N.
 XX (TAON/) TAO N.
 XX (WARR/) WARREN W C.
 XX Byatt JC, Mathialagan N, Tao N, Warren WC;
 XX WPI; 2003-110599/10.
 XX New nucleic acid associated with lactation, and muscle and fat
 XX deposition, useful for genome mapping, gene identification and analysis,
 XX cattle breeding, or for genetically improving cattle.
 XX Claim 2; SEQ ID NO 13579; 245pp; English.

The invention relates to a purified nucleic acid molecule associated with
 lactation or muscle and fat deposition (designated LMFD), derived from
 cattle, and the LMFD nucleic acid can specifically hybridise to a second
 nucleic acid molecule comprising any of 15112 nucleotide sequences,
 appearing as ABX34836-ABX49947, or complements of them. Also included are
 : (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
 acid linked to a promoter and a 3' non-translated sequence that
 functions in the cell to cause termination of transcription and addition
 of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 (2) determining a level or pattern of a molecule in a bovine cell or
 tissue comprising: (a) incubating a marker nucleic acid (comprising any
 of the 15112 nucleic acid sequences or its complement or fragment) with a
 complementary nucleic acid molecule obtained from the bovine cell or
 tissue, where hybridisation between the marker nucleic acid and the
 complementary nucleic acid permits the detection of the molecule; and (b)
 detecting the level or pattern of the complementary nucleic acid, where
 the detection of the complementary nucleic acid is predictive of the
 level or pattern of the molecule. The LMFD nucleic acid is used for
 determining a level or pattern of a molecule in a bovine cell or tissue.
 It is useful for genome mapping, gene identification and analysis, cattle
 breeding, preparation of constructs for use in cattle gene expression, or
 for genetically improving cattle. The present sequence is one of the
 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
 present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139

Sequence 432 BP; 91 A; 102 C; 125 G; 114 T; 0 U; 0 Other;
 Query Match 24.9%; Score 405.6; DB 8; Length 432;
 Best Local Similarity 97.7%; Pred. No. 2e-106;
 Matches 422; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 1083 GGGTGTGTGACAGGAGGATATCGAATGTTGAGATTTGCACTTCAAGTTGCTTTTAA 1142
 DB |||||
 QY 1143 GCCGACACCATCTATCGGGGTGAACACACACCGTGTCAAGGGAGCGTCAAGCTTGA 1202
 DB |||||
 QY 61 GCCGACACCATCTATCGGGGTGAACACACACCGTGTGTCAAGGGAGCGTCAAGCTTGA 120
 QY 1203 GCTTCTAGCAAGAGGGCGCCATGACCATGCTGCCCTCGAGCTGTTCCTGTGTTGA 1262

Db 121 |GCTTCTGCGAAGAGGCGCCATGACCCATGGTGGTCCCTCGAGCTGTTCTGTGGTGA 180
Qy 1263 ATCCATGCCCGGTTGGTCTCTCATGGACCAGCTGATGGCGACGTCGTGCTCAGTGGAGAT 1322
Db 181 ATCCATGCCCGGTTGGTCTCTCGTGACCGACGCTGATGGCGACGTCGTGCTCAGTGGAGAT 240
Qy 1323 GTTCGGCTCAATACTGCACTTCAAGAACCGAGTTGGCTCTTTCTAGCAGAGGCGAGGAC 1382
Db 241 GTTCGGCTCAATGCTGCACTTCAAGAACCGAGTTGGCTCTTTCTAGCAGAGGCGAGGAC 300
Qy 1383 ACCTGATGAGCTCGGCCCAATTTTATCATATTTATCATAGTAAATAGTAGCTCAAGCGTGG 1442
Db 301 ACCTGATGAGCTCGGCCCAATTTTATCATATTTATCATAGTAAATAGTAGCTCAAGCGTGG 360
Qy 1443 CTTGGTTGCTGCTCTTTCGACCGTAGTGTGTTTTTTTTTTT-TTCCCGCAAGTGTGATGC 1501
Db 361 CTTGGTTGCTGCTCTTTCGACCGTAGTGTGTTTTTTTTTTTCCCGCGCAAGTGTGATGC 420
Qy 1502 GATGAAGTGAAT 1513
Db 421 GATGAAGTGAAT 432

RESULT 9
ADT45459
ID ADT45459 standard; cDNA; 1086 BP.
AC ADT45459;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polynucleotide #20210.
XX

Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.

OS Bacteria.
XX
XX US2003233675-A1.
PN
XX
PD 18-DEC-2003.
PF
XX 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.

PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.

XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX

XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 43897; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a

CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX
SQ Sequence 1086 BP; 319 A; 201 C; 290 G; 276 T; 0 U; 0 Other;
Query Match 22.2%; Score 360.6; DB 13; Length 1086;
Best Local Similarity 58.9%; Pred. No. 3.4e-93;
Matches 643; Conservative 0; Mismatches 439; Indels 9; Gaps 1;

Qy 235 GGAACAACGTTTGGAAACTACTTTCAGGTTGGACCTATGTTGTAATCTCATGGGGTGGT 294
Db 4 GGCAATATTTTGGTCATTTATTTTCGCATTAGTACTTTTGGCGAGTCTCAGCGCGGGT 63
Qy 295 GTTGGTTGTGTATCATGTTGTGTCCACCTAGATTCCACTCTAGAGGAGACCTACAA 354
Db 64 GTGGGGGTTGTGATTTGTTGTTCTCCACAACTAGAAATTTCCGAGAGAATAATTCAG 123
Qy 355 GTTGAATTCGATCGAAGACGCGCCGCGACAGACAGAGATAAACCCTCCACAAGAGAGAGACT 414
Db 124 GTAGAACTAGATAGAGGCGCGCGGACAAAGTAAATTTACGACTCCCGCCAAAGAGCT 183
Qy 415 GATACATGCAAAATTTCTGTCAGGGACACATGAAGGGGTGACTCTGGAACGCCAATTTCTT 474
Db 184 GATACCTCGAGATTTATCTCAGGAGTATTTTGAAGGCAAAACGCTGGGAACCCCTATAACG 243
Qy 475 GTTATTCCTCCAAACACAGATCAATAGGAGTGATCACCCTGAAATAGCAATGTGTAC 534
Db 244 ATTTGTGATGTAATCAAGATACCTGTCGCCCAAGATTTAGCAGAGTGGCACAGATAT 303
Qy 535 CGACCTTCTCATGCAGACGCAACTTATGACTTCAAGTACGGTGTGTAGAGCTGTACAGGA 594
Db 304 CGGCTTCTCAGCGGATGCAACCTATGATGCAAAATATGGCATTCGCAATTTGGCAAGGT 363
Qy 595 GGTGGAGGTTCTCGGCGAGAAAACCGTTGGAAGGGTGGCTGCAGGGGCGCTCCCAAG 654
Db 364 GGGGTTAGTGGTCTGAGCAGCAATCGGAGAGTAGCAGCAGGTGGCATCGCTAAA 423
Qy 655 AAAATTCCTTAAGCTCAAAATGTGGATTAGAGATCTTGTGCTTTTGTTCCAAAGTGCATCAG 714
Db 424 AAAATTC-----CGTCAAGTTGCCAATGTCGAAATATTGCTTACGTTAAGCGG 474
Qy 715 GTTGTGCTCCAGAGACGCGGTTGATTTATGGGCTCTGTAACCTTTGGAAACAGATAGAGAGC 774
Db 475 ATTAAGACTTTGGAAGGTGTAGTCGATCCTTAATACTGTAACCTTTAGAACAAGTAGAAGC 534
Qy 775 AACATCGTTAGATGCTCTGATCCAGATACGCAGAGAGATGATAGCGCATCGACAGA 834
Db 535 AATATCGTTGCTGCTCCGATGCTGAATGTGGCATCGCATGATTTGAATTTGATTTGAGCAA 594
Qy 835 GTACGAGTTCGAGGGGATTCGGTGGTGGAGTGATCACATGCGTCGCTAGAAAAGCTTCCT 894
Db 595 ATAGGTAGACAAAGCGATTTCTATCGCGGTTGTCTAGATGTTGTGGCGGCAATGTACCG 654
Qy 895 CGCGGGTCCGGTTCTCTCTGTTTCACAAAGCTCGAATCCGAACCTGGCAAAAGCTATGCTT 954
Db 655 AAAGGTTTGGGCGAAACCAAGTATTTGATAAATTAGAAGCTGATATCGCTAAGGGTGTGATG 714

Db 595 ACAGTGACAAAGGTGATCTTATCGCGGTGTAGTGAATCGGTAGCGGAATGTTCT 654
Qy 895 CGCGGCTCGGTCTCTGTTGTCGAAGCTCGAATCCGAATCGGAATGTTGTTT 954
Db 655 AAAGCTTTGGGTGAACCTGTATTGATTAAGTTGGAAGCAGATATCGCCAAAGCTGTGATG 714
Qy 955 TCTATTCCTCGGAGCAACGGGTTGAGATTGGCAGCGGATTCGCGGAGCCGACTTGACA 1014
Db 715 TCTTTACCTGCTAGCAAAAGGTTTGAATCGGTTTCAGGTTTTCGCGGACTCTCTGAACA 774
Qy 1015 GGAAGTGAGCATATGATGAGTTTATATGATAAGGCTGGAAGTGTGAGACACGGACT 1074
Db 775 GGAATTGAGCATACGACGATATATATGATGAACGGTGAGATTGCGACAGTAAT 834
Qy 1075 AATCGTCGGGTGTTGTCAGGGAGGATATCGAATGTTGAGATTGTCATCTCAAGTT 1134
Db 835 AACCGTTCCGGTGGGATTCGAAGGGGTTATTCGAACGGGAGAGAATATCATTTTGCAGTT 894
Qy 1135 GCTTTAAGCCGACACCATCTATCGGGGTGAACAGACACCGTGTCAAGGGAGGCTCAG 1194
Db 895 GCATTTAAACCCACAGCCACGATTAAGAAAGAGCAAAACAGTTACTCGTGAAGGGCAA 954
Qy 1195 AACGTTGAGCTTCTAGCAAGGGCGCCATGACCCATCGCTCGCCCTCGAGCTGTTCT 1254
Db 955 GAAACCTTATTAGCGCCAAAGGTAGACATGATCTCTGTGTTCTACCCGTGAGTCCCG 1014
Qy 1255 GTGTTGAATCAATGCGCGCTGTTGTTCTCATGACACGCTGATGGCGCAGTGGCTCAG 1314
Db 1015 ATGGTAGAAGCAATGTTGGCGTGTGTTACTGTCGATCATTTGTTACGCCATCATGACAG 1074
Qy 1315 TGCAGAGTGT 1325
Db 1075 TGTAAGGTGCT 1085

RESULT 11
ADS47934

ID ADS47934 standard; cDNA; 1089 BP.

XX AC ADS47934;

XX DT 02-DEC-2004 (first entry)

XX DE Bacterial polynucleotide #2677.

XX KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; bs.

XX OS Bacteria.

XX PN US200323675-A1.

XX PD 18-DEC-2003.

XX PF 20-FEB-2003; 2003US-00369493.

XX PR 21-FEB-2002; 2002US-0360039P.

XX PA (CAOY/) CAO Y.

XX PA (HINK/) HINKLE G J.

XX PA (SLAT/) SLATER S C.

XX PA (CHEN/) CHEN X.

XX PA (GOLD/) GOLDMAN B S.

XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX DR WPI; 2004-061375/06.

XX PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX PS Claim 1; SEQ ID NO 26364; 122pp; English.

XX CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 1089 BP; 278 A; 245 C; 314 G; 252 T; 0 U; 0 Other;

Query Match 21.3%; Score 346.2; DB 13; Length 1089;

Best Local Similarity 58.1%; Pred. No. 5e-89;

Matches 634; Conservative 0; Mismatches 448; Indels 9; Gaps 1;

Qy 235 GGAACACGTTTGGAACTACTTTTCAGGTTGCAACCTATGGTGAATCTCTGCGGTGGT 294
Db 4 GGAACACGTTTGGCAGTTTATTTCGCATTACCACTTCGCGTGAGTCCCATGGGGCGGC 63
Qy 295 GTTGGTTGTGTATCACTGCTGTTGTCACCTAGATAATTCACCTCACTGAGGAGACTACAA 354
Db 64 GTGGGGTAAATTATTGATGTTGCCCTCCCGATTGGAATTTCCCCAGAGGAATCCAA 123
Qy 355 GTTGAATCGATCGAAGACGGCCGCGAGACAGACAGAGATAACCTCCACAAGAGGAGACT 414
Db 124 GTAGACCTAGACGGCGCGCGCAGGACAGACAGCAAAATTACCACTTCGTAAGGAAGCA 183
Qy 415 GATACATGCAAAATTTCTGTCAGGACACATGAGGGGTGACTACTGGAACGCCAATTTCT 474
Db 184 GATCAATGCGAAATTTCTCTGCGGTGTTTCGAGGGTAAAACTCTGGGTACACCATCGCC 243
Qy 475 GTTATTGTCCAAACACAGATCAATAGGAGTGTATCACCGTGAATAGCCCAATGTGTAC 534
Db 244 ATTTGTGCGAATATAGGATGCGCGTCCAGAGATTACATGAGATGCGGTGAATAT 303
Qy 535 CGACCTTCTCATGCAGACGCAACTTATGACTTCAAGTACCGGTGTTAGAGCTGTACAGGA 594
Db 304 CGTCTTTCCACGCGCATGCCACCTATGAAGCGAAATATGGCATTCGCAATTTGGCAAGGA 363
Qy 595 GGTGGAGGTCTTCGGCAGAAAACCGTTGGAAGGGTGGCTGCAGGGCCCTCCCAAG 654
Db 364 GGAGGTGCTCTTCTGCTAGGGAACCATTTGGCGGGTGGCGCGAGGGCGATCGCCAAA 423
Qy 655 AAAATTCTTAAGCTCAAAATGTGAATTAGAGATCTTGTCTGTTTTCCTCAAGTGCATCAG 714
Db 424 AAGATTCTGGCCCAAGTTTAACGGGTGGAATTTGTGGCTATGTGAAAGATATCCAAGAC 483
Qy 715 GTTGTGCTCCCAAGACGGCGGTGATATATGGGTCTGTAACTTTGGAACAGATAGAGAC 774
Db 484 ATT-----GAGGCGACGGTGGACAGCAATACGGTGACCTTTGGAGCAAGTGAAGC 534

Db 362 GCAACCAAGGTTTGAATTTGGAGTGGATTTGCGAGTACTTTCTTGACCGGTAGTGAA 421
Qy 1024 CATATGATGAGTTTATATGATPAAGCTGGAAGTGTGAGACACGAGCTAAATCGCTCG 1083
Db 422 CATATGATGAGTTTATATGATGAACTGGAAGTGTGAGACACGAGCTAAATCGCTCG 481
Qy 1084 GGTGTGTGCGAGGAGGATATCGAATGTTGAGATTTGCACTTCAAAAGTTGCTTTTAAAG 1143
Db 482 GGTGGATACAGGGTGGAAATCCATGGGAAATTTAAATATAGAGTGTAGCTTTCAAG 541
Qy 1144 CCGACACCATCTATCGGGGTGAAACAGAACACCGGTGTCAAGGGAGCGTCAAGAGTTGAG 1203
Db 542 CCAACAGCTACAATTTGGTGAAGAAACAGCACACAGTGTCTCGAGAAAAGAG-AGATAGAA 600
Qy 1204 CTTCTAGCAAGAGGCGGCATGACCCATGCGTGTGCGCCCTCGAGCTGTTCTGTGGT 1259
Db 601 CTACTAGCCCGGGTGTGTCATGATCCTTGTGTGTGTCTGCGAGCGTCCGATGTT 656

RESULT 13

ACL25323/c
ID ACL25323 standard; DNA; 731 BP.

AC ACL25323;

XX 27-OCT-2003 (revised)

DT 17-OCT-2003 (first entry)

XX DNA clone originating in barley containing SNP encoding sequence #15314.

XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW gene; ss.

XX Hordeum vulgare; ssp. spontaneum.

XX WO2003057877-A1.

XX 17-JUL-2003.

XX 16-DEC-2002; 2002WO-IB005403.

XX 20-DEC-2001; 2001JP-00387059.

XX 20-DEC-2001; 2001JP-00387131.

XX 20-DEC-2001; 2001JP-00403299.

XX 20-DEC-2001; 2001JP-00403300.

XX 27-SEP-2002; 2002JP-00327515.

XX (UYNI-) UNIV JAPAN OKAYAMA.

XX Sato K, Takeda K, Kohara Y;

XX WPI; 2003-587127/55.

XX Single nucleotide polymorphism sites in barley varieties and DNA
PT sequences containing them for analysis and identification of barley
PT varieties and production of barley transformants with desired
PT characteristics.

XX Disclosure; SEQ ID XX; 284pp; Japanese.

XX The present invention relates to oligonucleotide clones originating in
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)

SQ Sequence 731 BP; 185 A; 177 C; 171 G; 195 T; 0 U; 3 Other;
Query Match 19.0%; Score 308.2; DB 9; Length 731;
Best Local Similarity 79.6%; Pred. No. 3.9e-78;
Matches 364; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
Qy 905 GTTCTCTGTGTCGACCAAGCTCGAATCGCAACTGGCAAAAGCTATGCTTTCTATTCTCTG 964
Db 731 GCTCTCTGTATTGACAAACTTTGAGCTCTACTTGGCAAGGCTATGCTTTCTCTCTG 672
Qy 965 CGAGCAACGGGTTTCGAGATTGGCAGCGGATTCGCGGGACCGACTTGCAGAGGAAGTGAGC 1024
Db 671 CAAGCAAGGGGTTTGAGATCGGTAGTGGATTTGCAGGTACTGACCTAACTGGAAGTGAGC 612
Qy 1025 ATATGATGATTTTATATGATTAAGCTGGAAGTGTGAGACACGAGCTAAATCGCTCGG 1084
Db 611 ATAACGATGAGTTCTATATGATGAGGTGGAATGTGAGAACACACGAACCAATCGCTCGG 552
Qy 1085 GTGTGTGTCAGGGAGGATATCGAATGTTTCAGATTGTGCACCTTCAAAAGTTGCTTTTAAAG 1144
Db 551 GCGGTGTACAGGGCGGATATCAATGGTGAACCATATATCTTCAAAAGTAGCTTTCAAGC 492
Qy 1145 CGACACCATCTATCGGGGTGAAACAGAACACCGGTGTCAAGGGAGCGTCAAGAGCTTGAGC 1204
Db 491 CAACAGCAACTATTGGGAAGAAAGCAAAATACTGTGACAAGGGATCATGAGGATATCGAGC 432
Qy 1205 TTCTAGCAAGAGGCGGCATGACCCATGCGTGTGCGCCCTCGAGCTGTTCTGTGGTGAAT 1264
Db 431 TTCTGACAAGGGGTGCGCATGACCCATGTGTGTCTCGGGCTGTTTCCCAATGGTGGAGA 372
Qy 1265 CCATGGCGGTTGGTCTCTCATGACACCGCTGTATGGGCGCACGTGGCTCAGTGCAGATGT 1324
Db 371 CGATGGCGGATTGGTCTCTCATGACACCGCTGTATGGGCGCACATGTTGCTCAATGCGAGATGT 312
Qy 1325 TCGCGCTCAATATCTGCACTTCAAGAACCAAGTTGCGCTC 1361
Db 311 TCCGCTGAACCTCGCCCTACAAGAACCAATCGGTT 275

RESULT 14

ACL25321/c

ID ACL25321 standard; DNA; 699 BP.

AC ACL25321;

XX 27-OCT-2003 (revised)

DT 17-OCT-2003 (first entry)

XX DNA clone originating in barley containing SNP encoding sequence #15312.

XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW gene; ss.

XX Hordeum vulgare; var. (cul.Haruna Nijo).

XX WO2003057877-A1.

XX 17-JUL-2003.

XX 16-DEC-2002; 2002WO-IB005403.

XX 20-DEC-2001; 2001JP-00387059.

XX 20-DEC-2001; 2001JP-00387131.

XX 20-DEC-2001; 2001JP-00403299.

XX 20-DEC-2001; 2001JP-00403300.

XX 27-SEP-2002; 2002JP-00327515.

XX (UYNI-) UNIV JAPAN OKAYAMA.

XX Sato K, Takeda K, Kohara Y;

XX WPI; 2003-587127/55.

Search completed: August 27, 2005, 01:03:33
Job time : 875 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 21:25:10 ; Search time 4902 Seconds
(without alignments)
12625.969 Million cell updates/sec

Title: US-10-677-179-7

Perfect score: 1626

Sequence: 1 gcacgagctcagcttcgtct.....caaaaaaaaaaaaaaaaaaaaa 1626

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsl:*
9: gb_gsl2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1626	100.0	1626	AY105171	Zea mays
C 2	618.2	38.0	657	BM073857	MEST118-F
C 3	607.2	37.3	620	BM078362	CK210372
4	562	34.6	1160	7	CB654761
5	550.2	33.8	763	6	OSJNEC071
6	527.6	32.4	1488	3	CNSOACW
7	523.6	32.2	1494	3	CNSOACW
8	513	31.5	814	7	CN144654
9	499.8	30.7	1538	3	CNSOACW
10	498.6	30.7	723	7	CF303160
11	487.6	30.0	710	4	BU099249
12	480.4	29.5	702	5	BU099249
13	477.4	29.4	864	6	CA189902
14	475.6	29.2	694	6	CD870292
15	473	29.1	690	6	CA124299
16	471.2	29.0	696	5	BQ841547
17	469.2	28.9	667	6	CA119896
18	468.2	28.8	749	7	CA094750
19	461	28.4	812	7	CN129049
20	458.8	28.2	689	6	CA298301
21	457.8	28.2	673	7	CF571696
22	455.2	28.0	771	6	CA102511
23	449.6	27.7	923	7	CO082566
24	443.2	27.3	718	5	BU100489

25	443	27.2	639	4	BG414940	HVSMEK000
26	442.2	27.2	729	5	BQ789431	WHE4161_C
27	435	26.8	605	6	CA220612	SCRFL400
28	433.4	26.7	629	1	AV832717	AV832717
29	426.4	26.2	661	6	CD899216	CD899216
30	425.4	26.2	604	6	CA214673	CA214673
31	419.6	25.8	644	1	AJ432117	AJ432117
32	419.2	25.8	579	6	CA281236	SCRFL105
C 33	419	25.8	596	4	BG840749	CA281236
34	411.4	25.3	683	6	CD207597	MEST11-D0
C 35	409.2	25.2	489	4	BM350343	CD207597
36	408	25.1	697	6	CD890775	HSL_33_DI
37	408	25.1	865	7	CO114396	CD890775
38	407.6	25.1	568	2	BF065345	G118_115G
39	405.8	25.0	592	6	CD868036	CO114396
40	398.6	24.5	835	7	CO080629	BF065345
41	398.4	24.5	666	7	CK369578	HV_Eb001
42	396.2	24.4	562	1	AJ435616	CK369578
43	395.2	24.3	604	2	BE497989	AJ435616
44	394.8	24.3	755	7	CK279831	BE497989
45	390.8	24.0	786	4	BG600177	CK279831

ALIGNMENTS

RESULT 1	AY105171	Zea mays	1626 bp	linear	HTC 16-OCT-2002
LOCUS	AY105171	Zea mays	1626 bp	linear	HTC 16-OCT-2002
DEFINITION	Zea mays	PCO087106 mRNA sequence.			
ACCESSION	AY105171				
VERSION	AY105171.1	GI:21208249			
KEYWORDS	HTC.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
REFERENCE	1 (bases 1 to 1626)				
AUTHORS	Hayney,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.				
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes				
JOURNAL	Unpublished (2002)				
REFERENCE	2 (bases 1 to 1626)				
AUTHORS	Coe,B.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA				
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSU, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.				
FEATURES	location/Qualifiers				
source	1..1626				
	/organism="Zea mays"				
	/mol_type="mRNA"				
	/db_xref="MaizeDB:634903"				
	/db_xref="taxon:4577"				
	/clone_lib="Maize Mapping Project/DuPont Consensus Library"				
	/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"				
ORIGIN					
Query Match	100.0%	Score	1626	DB 3	Length 1626;

Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	GCACGAGCTCAGCTTCGTCTCTCTCGCGCGCGCGCAGCAACTATCATCATCTTCATTAG	60
Db	1	GCACGAGCTCAGCTTCGTCTCTCTCGCGCGCGCGCAGCAACTATCATCATCTTCATTAG	60
Qy	61	CTCATCTCAATCTATTCCGATGACGAGCTGCCAGGCCACAGCGAGTGGCCCACTCAGG	120
Db	61	CTCATCTCAATCTATTCCGATGACGAGCTGCCAGGCCACAGCGAGTGGCCCACTCAGG	120
Qy	121	GCACGGCTCGCACCCCGCGGCGATCGCGGCTTGTCTGAGATTTCGCCAGCCCTCTCTCC	180
Db	121	GCACGGCTCGCACCCCGCGGCGATCGCGGCTTGTCTGAGATTTCGCCAGCCCTCTCTCC	180
Qy	181	CTCCGCTTCGCGCTGCACCGCTGCCGACCTGCTCGCTTAGAGGTGAAGGCAATCTGGAAC	240
Db	181	CTCCGCTTCGCGCTGCACCGCTGCCGACCTGCTCGCTTAGAGGTGAAGGCAATCTGGAAC	240
Qy	241	ACGTTTGGAACTACTTTTCAGTTGCAACCTATGCTGAATCTCATGGGGGTGGTGGT	300
Db	241	ACGTTTGGAACTACTTTTCAGTTGCAACCTATGCTGAATCTCATGGGGGTGGTGGT	300
Qy	301	TGTGTTATCAGTGGTGTCCACTAGAAATTCACCTCACTGAGGCGAGACTTACAAGTTGAA	360
Db	301	TGTGTTATCAGTGGTGTCCACTAGAAATTCACCTCACTGAGGCGAGACTTACAAGTTGAA	360
Qy	361	CTCGATCGAAGACGCGCGCGCAGAGCAGATAAATCCTCCACAAGGAAGGAGCTGATACA	420
Db	361	CTCGATCGAAGACGCGCGCGCAGAGCAGATAAATCCTCCACAAGGAAGGAGCTGATACA	420
Qy	421	TGCAAAATCTGTGAGGAGCAGATGAGGGGTGACTCTGGAACGCCAAATCTTGTATT	480
Db	421	TGCAAAATCTGTGAGGAGCAGATGAGGGGTGACTCTGGAACGCCAAATCTTGTATT	480
Qy	481	GTCCCAAAACACAGATCAATAGGCACTGATCACCGTGAAATAGCCAAATGTGTACCGACT	540
Db	481	GTCCCAAAACACAGATCAATAGGCACTGATCACCGTGAAATAGCCAAATGTGTACCGACT	540
Qy	541	TCTCATGACAGCGAACTTATGACTTCAAGTACGGTGTAGAGCTGTACAGGAGGTGG	600
Db	541	TCTCATGACAGCGAACTTATGACTTCAAGTACGGTGTAGAGCTGTACAGGAGGTGG	600
Qy	601	AGGTCCTCGGCGAGAAAACGTTGGAAGGGTGGCTGAGGGCCCTCCCAAGAAAT	660
Db	601	AGGTCCTCGGCGAGAAAACGTTGGAAGGGTGGCTGAGGGCCCTCCCAAGAAAT	660
Qy	661	CTTAAGCTCAATGTGGATTAGAGATCTTGTCTTGTTCCTCAAGTGCATCAGGTGTG	720
Db	661	CTTAAGCTCAATGTGGATTAGAGATCTTGTCTTGTTCCTCAAGTGCATCAGGTGTG	720
Qy	721	CTCCAGAGAGCGGGTGTATATGGGTCTGTAACTTTGGAAACAGATAGAGAGCAATC	780
Db	721	CTCCAGAGAGCGGGTGTATATGGGTCTGTAACTTTGGAAACAGATAGAGAGCAATC	780
Qy	781	GTTAGATGCTCTGATCCAGAGTACCCAGAGAGATGATAGCGCAATCGACAGAGTACGA	840
Db	781	GTTAGATGCTCTGATCCAGAGTACCCAGAGAGATGATAGCGCAATCGACAGAGTACGA	840
Qy	841	GTTTCAGGGGATTCGGTTCGGTGGAGTGATCATCGTTCGCTAGAAACGTTCTCGCGGG	900
Db	841	GTTTCAGGGGATTCGGTTCGGTGGAGTGATCATCGTTCGCTAGAAACGTTCTCGCGGG	900
Qy	901	CTCGGTTCTCTGTGTTTCGACAGCTCGAATCCGAACTGGCAAAAGCTATGCTTTATT	960
Db	901	CTCGGTTCTCTGTGTTTCGACAGCTCGAATCCGAACTGGCAAAAGCTATGCTTTATT	960
Qy	961	CTCGGAGCAACGGGTTTCGAGATTTCGAGCGGATTCGCGGGACCGACTTCACAGAGT	1020
Db	961	CTCGGAGCAACGGGTTTCGAGATTTCGAGCGGATTCGCGGGACCGACTTCACAGAGT	1020
Qy	1021	GAGCAATAATGAGTGTATATATGATAAGGCTGGAAGTGTGAGGACCGGCAATATCGC	1080

Db	1021	GAGCAATAATGATGAGTTTTATATGATAAGGCTGGAAGTGTGAGGACACGAGCTAATCGC	1080
Qy	1081	TCGGTGGTGTGACAGGAGGGATATCGAATGTTGAGATTGTGCACITTCAAAGTTGCTTTT	1140
Db	1081	TCGGTGGTGTGACAGGAGGGATATCGAATGTTGAGATTGTGCACITTCAAAGTTGCTTTT	1140
Qy	1141	AAGCCGACACATCTATCGGGGTGAAACAGAACACCGTGTCAAGGGAGCGTCAGAACGTT	1200
Db	1141	AAGCCGACACATCTATCGGGGTGAAACAGAACACCGTGTCAAGGGAGCGTCAGAACGTT	1200
Qy	1201	GAGCTTCTAGCAAGAGGGCGCCATGACCCATGCGTCCGCTCGAGCTGTTCTCTGGTG	1260
Db	1201	GAGCTTCTAGCAAGAGGGCGCCATGACCCATGCGTCCGCTCGAGCTGTTCTCTGGTG	1260
Qy	1261	GAATCCATGCGCGGTGGTCTCTCATGGACAGCTGATGGCGCACGTTGGCTCAGTGGAG	1320
Db	1261	GAATCCATGCGCGGTGGTCTCTCATGGACAGCTGATGGCGCACGTTGGCTCAGTGGAG	1320
Qy	1321	ATGTTTCGCGCTCAATACTGCACCTTCAAGAACCAAGTTGGCTCTTTCTAGCAGAGCAGC	1380
Db	1321	ATGTTTCGCGCTCAATACTGCACCTTCAAGAACCAAGTTGGCTCTTTCTAGCAGAGCAGC	1380
Qy	1381	ACACCTGATGAGCTCGCGCCAAATTTTATCATTTATCATAGTAATAGTAGCTCAAGCGT	1440
Db	1381	ACACCTGATGAGCTCGCGCCAAATTTTATCATTTATCATAGTAATAGTAGCTCAAGCGT	1440
Qy	1441	GGCTTGGTGTGCTGTCTCTTGACCGTAGTTTGTGTTTTTTTTTTTTTCCGCAAGTGTATG	1500
Db	1441	GGCTTGGTGTGCTGTCTCTTGACCGTAGTTTGTGTTTTTTTTTTTTTCCGCAAGTGTATG	1500
Qy	1501	CGATGAAGTGAATAAGGCACTTGGTTTCCTGTGTCATTGTACACGTTTTCATATAATGTAA	1560
Db	1501	CGATGAAGTGAATAAGGCACTTGGTTTCCTGTGTCATTGTACACGTTTTCATATAATGTAA	1560
Qy	1561	TTACTTTCGAAGATGATGCAATTTTATAGATGTGGCTTGTGAAGACAAAAA	1620
Db	1561	TTACTTTCGAAGATGATGCAATTTTATAGATGTGGCTTGTGAAGACAAAAA	1620
Qy	1621	AAAAA 1626	
Db	1621	AAAAA 1626	
RESULT 2			
BM073857/c			
LOCUS	BM073857	657 bp	mrna linear EST 13-NOV-2001
DEFINITION	MEST76-G12.T3 ISUM4-TN Zea mays cdna clone MEST76-G12.3', mrna		
ACCESSION	BM073857		
VERSION	BM073857.1		
KEYWORDS	EST.		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
AUTHORS	Qiu, F., Cui, F., Guo, L., Ashlock, D.A., Wen, T.J. and Schnable, P.S.		
TITLE	Expressed Sequence Tags from B73 Maize Seedlings and Silks		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Patrick S. Schnable Schnable Laboratory Iowa State University G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA Tel: 515-294-0975 Fax: 515-294-2299 Email: schnable@iastate.edu Individual basecall and confidence value were assigned using the Phred software, (http://depts.washington.edu/ventures/collabtr/direct/index.htm#b rt). Overall sequence quality assessment and vector trimming were conducted using the Lucy software (http://www.tigr.org/softlab/lucy). Lucy parameters were set to ensure an overall trimmed quality of		

97.5% or better without any vector fragments in the chosen high-quality region of each sequence. Low-quality bases between the poly-T and the high-quality region were replaced with N's to serve as spacers.

PCR Primers

FORWARD: Forward PCR primer sequence, primer T7-1 (AA TAC GAC TCA CTA TAG)
 BACKWARD: Backward PCR primer sequence, primer T3 (ATT AAC CCT CAC TAA AG)

Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).

FEATURES

source

Location/Qualifiers

1..657

/organism="Zea mays"

/mol_type="mRNA"

/cultivar="B73"

/db_xref="taxon:4577"

/clone="WEST76-G12"

/tissue_type="Seedling and silk"

/lab_host="DH10B"

/clone_lib="ISUM4-TN"

/note="Vector: pT7T3PAC; Site 1: EcoRI; Site 2: NotI; ds-cDNA molecules were generated as follows. First-strand cDNA was prepared from oligo-dT selected mRNA by priming with a NotI oligo-dT primer (5' AACTGGAAGATTCGCCCGCAGAAATTTTTTTTTT). The resulting DNA:RNA hybrid was treated with RNase H and used as a template for DNA PolI-catalyzed second strand synthesis. After the addition of EcoRI adaptors, the ds-cDNAs were digested with NotI and size-selected. The resulting molecules were directionally cloned into the EcoRI and NotI sites of the pT7T3PAC vector. The library then went through one round of normalization to Cot value of 5 based on the methods of Marcelo Bento Soares (Genome Research 6: 791-806, 1996)."

ORIGIN

Query Match

Best Local Similarity 98.5%; Score 618.2; DB 4; Length 657;

Matches 645; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

QY 973 GGTTTCAGATTGGCAGCGGATTCGCGGACCGACTTCACAGGAGTGAGCATATGAT 1032
 DB 657 GGTTTCAGATTGGCAGCGGATTCGCGGACCGACTTCACAGGAGTGAGCATATGAT 598
 QY 1033 GAGTTTATATGATGAAGCTGGAAGTGTGAGATTCGAGATTAATCGCTCGGTTGTG 1092
 DB 597 GAGTTTATATGATGAAGCTGGAAGTGTGAGATTCGAGATTAATCGCTCGGTTGTG 538
 QY 1093 CAGGGAGGATTCGAATGTTGAGATTGTGCACTTCAAAAGTTGCTTTAAAGCCACCA 1152
 DB 537 CAGGGAGGATTCGAATGTTGAGATTGTGCACTTCAAAAGTTGCTTTAAAGCCACCA 478
 QY 1153 TCTATCGGGGTGAACAGACACCGTGTCAAGGAGCGTCAGACGTTGAGCTTCTAGCA 1212
 DB 477 TCTATCGGGGTGAACAGACACCGTGTCAAGGAGCGTCAGACGTTGAGCTTCTAGCA 418
 QY 1213 AGAGGCGCCATGACCCATGCTGCCCTTCGAGCTGTTCTGTGTTGGAATCATGGCC 1272
 DB 417 AGAGGCGCCATGACCCATGCTGCCCTTCGAGCTGTTCTGTGTTGGAATCATGGCC 358
 QY 1273 GGTGTGTCCTATGACACAGTGAATGCGGCACGTGGCTCAGTGCAGAGATGTTCCGCTC 1332
 DB 357 GGTGTGTCCTATGACACAGTGAATGCGGCACGTGGCTCAGTGCAGAGATGTTCCGCTC 298
 QY 1333 AATAGTCATCTCAAGAACCAAGTTGGCTTCTTCTAGCAGAGGCGAGCACACCTGATGAG 1392
 DB 297 AATAGTCATCTCAAGAACCAAGTTGGCTTCTTCTAGCAGAGGCGAGCACACCTGATGAG 238
 QY 1393 CTCGGCCCAATTTATCATTTATCATAGTAGTCTCAAGCGTGGCTGGTTTCG 1452
 DB 237 CTCGGCCCAATTTATCATTTATCATAGTAGTCTCAAGCGTGGCTGGTTTCG 178
 QY 1453 TTGTCTCTGACCGTAGTTTGTGTTTTTTTTTCCCGCAGTGTGATGCGATGAAGTGA 1512

Db 177 TTGTCTCTTCCCGTAGTTTGTGTTTTTTTTTCCCGCAAGTGTGATGCGATGAAGTGA 118
 QY 1513 TAAGGCACCTGGTTTCTCTGTCATTTGTACAGTTTCATATAATG-TAATCTACTTCGAA 1571
 DB 117 TAAGGCACCTGGTTTCTCTGTCATTTGTACAGTTTCATATAATGCTAATCCCTTCGAA 58
 QY 1572 GATGATGCAATTTTATAGATGGCTGTGAAGACAAAAAATAAAAAA 1626
 DB 57 GATGATGCAATTTTATAGATG-GGCTTGTGAAGCAAAAAAATAAAAAA 4

RESULT 3

BM078362/c

LOCUS

DEFINITION

BM078362

VERSION

BM078362.1

KEYWORDS

SOURCE

ORGANISM

Zea mays

Zea mays

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM078362 620 bp mRNA linear EST 14-NOV-2001
 MEST118-F09.T3 ISUM4-TN Zea mays cDNA clone MEST118-F09 3', mRNA
 sequence.

BM078362

BM078362.1

GI:16925294

EST.

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

177 TTGTCTCTTCCCGTAGTTTGTGTTTTTTTTTCCCGCAAGTGTGATGCGATGAAGTGA 118
 1513 TAAGGCACCTGGTTTCTCTGTCATTTGTACAGTTTCATATAATG-TAATCTACTTCGAA 1571
 117 TAAGGCACCTGGTTTCTCTGTCATTTGTACAGTTTCATATAATGCTAATCCCTTCGAA 58
 1572 GATGATGCAATTTTATAGATGGCTGTGAAGACAAAAAATAAAAAA 1626
 57 GATGATGCAATTTTATAGATG-GGCTTGTGAAGCAAAAAAATAAAAAA 4

BM078362 620 bp mRNA linear EST 14-NOV-2001
 MEST118-F09.T3 ISUM4-TN Zea mays cDNA clone MEST118-F09 3', mRNA
 sequence.

BM078362

BM078362.1

GI:16925294

EST.

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Zea mays

Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).
 Location/Qualifiers
 1..620

/organism="Zea mays"

/mol_type="mRNA"

/cultivar="B73"

/db_xref="taxon:4577"

/clone="MEST118-F09"

/tissue_type="Seedling and silk"

/lab_host="DH10B"

/clone_lib="ISUM4-TN"

/note="Vector: pT7T3PAC; Site 1: EcoRI; Site 2: NotI;

ds-cDNA molecules were generated as follows. First-strand

cDNA was prepared from oligo-dT selected mRNA by priming

with a NotI oligo-dT primer (5' AACTGGAAGATTCGCCCGCAGAAATTTTTTTTTT). The

resulting DNA:RNA hybrid was treated with RNase H and used

as a template for DNA PolI-catalyzed second strand

synthesis. After the addition of EcoRI adaptors, the

ds-cDNAs were digested with NotI and size-selected. The

resulting molecules were directionally cloned into the

EcoRI and NotI sites of the pT7T3PAC vector. The library

then went through one round of normalization to Cof value of 5 based on the methods of Marcelo Bento Soares (Genome Research 6: 791-806, 1996)."

ORIGIN	Query Match 37.3%; Score 607.2; DB 4; Length 620;
	Best Local Similarity 98.7%; Pred. No. 2.4e-165;
	Matches 612; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy	1002 GACCGACTTGACAGGAAGTGAACATATGATGAGTTTATATGATAGGCTGGAAGTGT 1061
Db	620 GACCGACTTGACAGGAAGTGAACATATGATGAGTTTATATGATAGGCTGGAAGTGT 561
Qy	1062 CAGGACAGGACTAATCGCTCGGTGGTGTCCAGGAGGGATATCGATGTTGATGTT 1121
Db	560 CAGGACAGGACTAATCGCTCGGTGGTGTCCAGGAGGGATATCGATGTTGATGTT 501
Qy	1122 GCACCTTCAAAAGTGTCTTTAAGCCGACACCATCTATCGGGTGAAACAGAACACCGTGT 1181
Db	500 GCACCTTCAAAAGTGTCTTTAAGCCGACACCATCTATCGGGTGAAACAGAACACCGTGT 441
Qy	1182 AAGGAGCGCTCAGAACGTTGAGCTTTAGCAAGAGGGCGCCATGACCCATCGTGGCCCC 1241
Db	440 AAGGAGCGCTCAGAACGTTGAGCTTTAGCAAGAGGGCGCCATGACCCATCGTGGCCCC 381
Qy	1242 TCGAGCTGCTCCTGTGGTGAATCCATGCGCGCTTGGTCTCATGGACAGCTGATGGC 1301
Db	380 TCGAGCTGCTCCTGTGGTGAATCCATGCGCGCTTGGTCTCATGGACAGCTGATGGC 321
Qy	1302 GCAGTGGCTCAGTGCAGATGTTGCGGCTCAATCTGCACTTCAAGAACAGTTGGCTC 1361
Db	320 GCAGTGGCTCAGTGCAGATGTTGCGGCTCAATCTGCACTTCAAGAACAGTTGGCTC 261
Qy	1362 TTTTACAGAGGCGAGACACACCTGATGAGCTCGCGCCAAATTTATCATTTATCATAG 1421
Db	260 TTTTACAGAGGCGAGACACACCTGATGAGCTCGCGCCAAATTTATCATTTATCATAG 201
Qy	1422 TAATTAAGTAGCTCAAGCGTGGCTGTTGTTGCTCTCTCTGACCGTAGTTTCTTTTT 1481
Db	200 TAATTAAGTAGCTCAAGCGTGGCTGTTGTTGCTCTCTCTGACCGTAGTTTCTTTTT 141
Qy	1482 TTTTCCCGCAAGTGTGATGCGATGAAGTGAATAAGGCACTTGGTTCTGTCGATTTGTA 1541
Db	140 TTTTCCCGCAAGTGTGATGCGATGAAGTGAATAAGGCACTTGGTTCTGTCGATTTGTA 81
Qy	1542 CAGCTTTCATATATGATTAATCTACTTCGAAGATGATGATTTTATAGATGCGTGTG 1601
Db	80 CAGCTTTCATATATGATTAATCTACTTCGAAGATGATGATTTTATAGATGCGTGTGTA 21
Qy	1602 AAAACAAAAAATAAAAAA 1621
Db	20 AAAAATAAAAAAATAAAAAA 1

RESULT 4	
CK210372	
LOCUS	CK210372
DEFINITION	FGAS022177 Triticum aestivum FGAS: Library 5 GATE 7 Triticum
	aestivum cDNA, mRNA sequence.
ACCESSION	CK210372
VERSION	CK210372.1
KEYWORDS	EST.
SOURCE	Triticum aestivum (bread wheat)
ORGANISM	Triticum aestivum
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
	Poideae; Triticeae; Triticum.
	1 (bases 1 to 1160)
	Allard, P., Crosby, W.L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D.,
	Genswein, B., Grat, R., Gulick, P., Hrycan, L.D., Laroche, A.,
	Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilsson, D.,
	Penniket, C., Roach, J.L. and Sarhan, F.
TITLE	Functional Genomics of Abiotic Stress In Wheat and Canola Crops

JOURNAL COMMENT

Unpublished (2003)
Contact: Wm L Crosby
Bioinformatics
University of Saskatchewan, Department of Computer Science
1C101 Engineering Building, 57 Campus Drive, Saskatoon,
Saskatchewan, S7N 5A9, Canada
Tel: 306 966 1769
Fax: 306 966 2033
Email: fgas_estcs.usask.ca
This sequence is the direct result of the Base calling software
Phred (default parameters). It is the raw base calls. To aid in the
identification of the high quality insert the software Lucy
(default parameters) has been run on this sequence. Lucy identified
the region [18,834].
Plate: USB022 row: K column: 15.

FEATURES source

1. 1160
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone_lib="Triticum aestivum FGAS: Library 5 GATE 7"
/notes="Vector: pCMV.SPORT6; Crown and developmental stages
of spike formation in wheat cultivar Norstar. 4 mRNA
populations were combined before constructing the library.
The first mRNA population is from lcn crown sections after
30 days of cold acclimation. The second is from lcn crown
sections after 11 days of deacclimation (before
deacclimation plants were fully vernalized for 49 days).
The third is from different developmental stages of spike
formation (5 to 50mm) that still have not emerged from the
leaf (dissection required). The last is from different
developmental stages of spike and seed formation after
having emerged from the leaf (visible). First strand
synthesis in this library was done in the presence of
methylated dCTP thereby protecting from internal cleavage
with NotI."

ORIGIN

Query Match 34.6%; Score 562; DB 7; Length 1160;	
Best Local Similarity 72.5%; Pred. No. 4.4e-152;	
Matches 810; Conservative 0; Mismatches 271; Indels 36; Gaps 5;	
Qy	25 CGCGCGCGCGGAGCAATCATCATCTTTCATTAGCTCATCAATCTATTCCGATGACG 84
Db	14 CGNGACCCAGCGGTGCGCGGAAACCAACCAACCAATCTACCAACCTCTCCAAATGGCA 73
Qy	85 ACCGTGCCCCAAGCACAGCAGGTGGGCGCATCTACGGGCGAGGTGCGACCGCGCGCATC 144
Db	74 ACCGCGCCACGTGCGACCAAGCTCGCGCGCGGCGGCGCTGGAGCTCCCTCCCGCGCGC 133
Qy	145 GCGCGCTTGTGGAGTTTGCGCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 204
Db	134 GGGTTCCGGGCGCTCAAGGACTCCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCT 193
Qy	205 CGCACTGCTCGCTAGAGGTGAAGGCATCTGGAAACACAGTTTGGAAATCTACTTTCAGGTT 264
Db	194 CGCGCGCGCGCTAGAGGTGAAGGCATCTCTGGGAGTACTTCTCCAGGTT 253
Qy	265 GCAACCTATGTTGAATCTCATGGGGGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 324
Db	254 GCAACTTATGGAGAGTCTCATGGAGGGGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGT 313
Qy	325 AGAATTCCTCACTGAGGCGAGGACCTACAGATTGAACTCGATCGAAGACGCGCGGACAG 384
Db	314 AGAATTCCTCACTGAGGAGGAGACATG-----CAG 343
Qy	385 AGCAGATAAATCTCCCAAGGAGGAGACTGATACATGCAAAATTTCTGTCCAGGACACAT 444
Db	344 AGCAGATAAATCTCCCAAGGAGGAGACCGATCTTGTAAATCTTTCCAGGACATAC 403
Qy	445 GAAGGGGTGATCTAGTGAACCGCAATTTCTTTGTTATTTGTTCCCAACACAGATCAAAATAGG 504
Db	404 GAAGGAATGACCACTGGGACACCAATTTCTTTGTTTCTTCTTCCCAACACAGGATCAAAAGGG 463

VERSION	BX818532.1	GI:42473287	
KEYWORDS	HTC; GSI.T.CDNA		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.		
TITLE	Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1494)		
AUTHORS	Genoscope.		
JOURNAL	Direct Submission		
COMMENT	Submitted (18-NOV-2003) Genoscope - Centre National de Sequençage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)		
	The sequences are based on single pass reads.		
	Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.		
	Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.		
	URGV INRA : Clepet C., Caboche M.		
	Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.		
	http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full length		
	http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.		
FEATURES	Location/Qualifiers		
source	1..1494		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/strain="Col-0"		
	/db_xref="taxon:3702"		
	/clone="GSLTSL1842B05"		
	/tissue_type="Silique"		
	/plasmid="pCWS3PORT 6"		
gene	complement(1..1494)		
	/gene="At1g48850"		
ORIGIN			
Query Match	32.2%; Score 523.6; DB 3; Length 1494;		
Best Local Similarity	66.7%; Pred. No. 7.5e-141;		
Matches	762; Conservative 0; Mismatches 379; Indels 1; Gaps 1;		
QY	216 CCTAGAGGTGAGGCGATCTGGAACACAGTTCGAACTACTTTCAGGTTGCAACCTATGG	275	
DB	210 CTTCCAGATACAGCTACTGGAAGTTTCATATGGGACTCAATTCGAGTTTCAACTTTGG	269	
QY	276 TGAATCTCATGGGGTGGTGTGTGTATCATGTTGTCACCTAGAAATCCACT	335	
DB	270 AGATACATGGAGGAGGTGGTGTATCATTTGATGTTGTCTCTCTGATTTCCACT	329	
QY	336 CACTGAGGAGCCTCAAGTTGAATCGATCGAAGACCGCCGGACAGACAGAGATAAC	395	
DB	330 TACTGAATCTGATTTGCAATTCGATCTCGATAGAGGCGCTGGTCAGAGCAGGATCAC	389	
QY	396 CTCACAGAGGAGAGATGATACATGCAAAATTCGTGAGGACACATGAGGGGTGAC	455	
DB	390 AACTCTAGAAAAGAGACTGATCTTGGCGGATATCGTCTGGAGTCTCTGAGGAGTATG	449	
QY	456 TACTGGAACCCCAATTTCTGTTATTCGCCAACAACAGATCAATCAAGCAGTCAACCG	515	
DB	450 GACAGGACACCTATTCATGTTGTGTATCAACAACAGATCAGAGGAGCTTGAATACAG	509	
QY	516 TGAATAGGCAATGTGTACCGACCTTCTCATGCGACGCAACTTATGACTTCAAGTACGG	575	

Db	510	TGAAATGTGCGTGTGCTTATAGACCATCGCATGCTGATGCACTTATGACATGAAGTATGG	569
QY	576	TGTTAGAGCTGTACAGGAGGTGGAGGTCTCTGGGCGAGAAACCCGTTGGAAGGTGGC	635
Db	570	TGTGAGATCATGTGAGGAGGTGGAGAGATCTTCAGCTAGAGAGACCATTGAAGAGTTC	629
QY	636	TGCAGGGGCCCTCCCAAGAAAAATTTCTTAAGCTCAAAATGTGGATTAGAGATCTTTGTCTT	695
Db	630	TCCTGGAGCTTTGGCCAAAGAAAAATTTGAAGCAATTTGCAGAACTGAGATCTTTCGCTA	689
QY	696	TGTTTCCAAAGTGCATCAGAGTTGTGCTCCCAAGAGACCGGTTGATTTATGAGTCTGTAAAC	755
Db	690	TGTCTCGCAATTCACCATGTTGTACTTCCAGAGAAATTTGTAGACCAACGAGAAATTAAC	749
QY	756	TTTGGAAACAGATACAGAGCAACATCGTTAGATGCTCTGATCCAGAGTACGACAGAAAGAT	815
Db	750	ACTGAAACAGATAGAAATAACATTTGTCAGATGCCCTAATCCCGAGTATGCGGAAAGAT	809
QY	816	GATAGAGCGCAATCGACAGAGTACGAGTTTCGAGGGGATTCGGTTCGGTGGAGTGCATCATG	875
Db	810	GATAGCTGGGATTCATGCTGTGTCAGACAAAAGGGAACCTCTGTTGGTGTGTGTGACCTG	869
QY	876	CGTCTGATAGAAAGTTCCTCGGGGCTCGGTTCTCTGTTGTTCCAGAGCTCGAATCCGA	935
Db	870	CATTGTTGGGAATGCTCCACGCTGGGCTTGGTACACCGGTTTTTCGATAAACTTGAAGCAGT	929
QY	936	ACTGCGAAAGCTATGCTTTCTTATTCCTGCGAGCAACGGGTTTCGAGATTTGGCAGCGATT	995
Db	930	ACTGCGAAAGCTTGTATGTCGCTACCTTGCACCTAAGGATTTGAGTTTGGAAACGGCTT	989
QY	996	CGCGGGACCCACTTTCAGCAGGAAGTGCAGCAATATGATGAGTTTATATGGAATAGGCTGG	1055
Db	990	TGCAGGTACCTTTTGGTCTTTGAACACAATGATGAGTTCTATATCCGATGAAAATGG	1049
QY	1056	AAGTGTGAGGACAGCAATATCGTCTGGTGGTGTGTCAGGAGGAGATATCGAATGTTGA	1115
Db	1050	AAGAATACGTACCAAGAACCAACCGATCTGTTGGAAATTCAGGAGGAGATCTCAAAATGGTGA	1109
QY	1116	GATTGTGCATCTCAAGTTGCTTTTAAGCCGACACCATCTATCGGGTGAACAGAACAC	1175
Db	1110	AATAATAACATGAGATGAGCTTTCAGCCAAACATCAACAATTCGAGGAGCAAAATATAC	1169
QY	1176	CGT-GTCAAGGGAGCGTCAGAACCGTTGAGCTTCTAGCAAGAGGGCGCCATGACCCATCG	1234
Db	1170	GGTAACCCAGACAGAGTAGAAGAACCGAAATGATTGGCGGTGGTCTCATGATCTCTGTG	1229
QY	1235	TCGCCCTTCGAGCTGTTCTGTTGGTAATCCATGGCGCGCTGGTCTCTCATGAGACAGC	1294
Db	1230	TTGTTCCACGAGCTGTGCCAATGGTGGAGCAATGGTGGCTCTAGTTCTTGTGGATCAAT	1289
QY	1295	TGATGGCGACGTGGCTCAGTGCAGATGTTTCGGCTCAATCTACTGCACTTCAAGAACAG	1354
Db	1290	TGATGGCGCAATACGCAATGCCATTTGTTTCCAAATAATCCAGAGTTGAGGAACTC	1349
QY	1355	TT 1356	
Db	1350	TT 1351	
RESULT 8			
LOCUS	CN144654		
DEFINITION	WOUND1_23_E02_g1_A002 Wounded leaves Sorghum bicolor cDNA clone		
ACCESSION	WOUND1_23_E02_A002 5', mRNA sequence.		
VERSION	CN144654.1		
KEYWORDS	EST.		
SOURCE	Sorghum bicolor (sorghum)		
ORGANISM	Sorghum bicolor		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.		
REFERENCE	1 (bases 1 to 814)		

AUTHORS	Cordonnier-Pratt, M.-M., Suzuki, Y., Sugano, S., Klein, R.R., Liang, C., Sun, F., Sullivan, R., Brady, J., Eastman, A., Miller, V., Gonzalez, M., Anfujo, C., Chhabra, D., Johnson, H., Kamran, D. and Pratt, L.H.
TITLE	A Sorghum EST database: mechanistically damaged and methyl jasmonate-treated leaves
JOURNAL	Unpublished (2003)
COMMENT	Contact: Cordonnier-Pratt MM Laboratory for Genomics and Bioinformatics The University of Georgia, Department of Plant Biology Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA Tel: 706 542 1860 Fax: 706 583 0210 Email: mmpratt@uga.edu Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in the Human Genome Center, University of Tokyo Institute of Medical Science; plant material and RNA prepared at Texas A & M University; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA. seq primer: Sug5 (CTTCTGCTCTAAAGCTGCG).
FEATURES	Location/Qualifiers 1..814 /organism="Sorghum bicolor" /mol_type="mRNA" /cultivar="BTx623" /db_xref="taxon:4558" /clone="WOUND1 23_E02 A002" /lab_host="DH10B-T1 phage-resistant E. coli" /clone_lib="Wounded leaves" /note="Organ: Leaf; Vector: pME18S-FL3; Site 1: XhoI; Site 2: XhoI; The library was prepared from polyA+ RNA harvested from 8-day-old hydroponically grown, BTx623 sorghum seedlings. For some plants, one-half of the second leaf was crushed without damaging the midvein. For others, methyl jasmonate was added to the growth medium to a final concentration of 100 uM. Leaves were harvested 3 and 27 hr after treatment and pooled. Double-stranded cDNA was of the cloned unidirectionally into different DraIII sites of the pME18S-FL3 vector (5-prime DraIII site is CACTGTGTG, 3-prime DraIII site is CACCATGTG). XhoI excises the cDNA insert."
ORIGIN	
Query Match	31.5%; Score 513; DB 7; Length 814;
Best Local Similarity	80.6%; Pred. NO. 8e-138;
Matches 613; Conservative	0; Mismatches 145; Indels 3; Gaps 1;
Qy	83 CGACCGTCCCAAGCCACAGCAGGTGGCGCACTCACGGGCGCGCTCGCACCCCGCGGA 142
Db	
Qy	57 CGCCCGTTCGCGAGCCGCGGTGGCCGCGCAAGGCGGTCCACGCGGTCTCTCCCGCGGGA 116
Db	
Qy	143 TCGCGCGCTTGTGAGTTTGCCCGAGCGCTCTCTCTCCCTCCGCTTCGCCGTGCACCGCT 202
Db	
Qy	117 TAGCGCGCTCCGAGTCCGCGCCCGCGTTCGCTTCGCTCGC---CGCGCC 173
Db	
Qy	203 GCGGCATCTGCTCGCTAGAGTGAAGCATCTTGAAACACCGTTTGGAACTACTTTTACG 262
Db	
Qy	174 GCGCGCGCTCGCTAGAGTGAAGGCATCGGGAATGTCTCGGGAATCTACTTCCAGG 233
Db	
Qy	263 TTGCAACTATGTGAATCTCATGGGGGTGTGTGGTTGTATCATAGTGGTTGTCAC 322
Db	
Qy	234 TTGCNACTATGGGGAATCCCATGGGGGGGTGTGGTTGTATCATAGTGGTGGCCAC 293
Db	
Qy	323 CTAGAAATCCACTCACTGAGGCAGACCTACAAGTTGAACTCGATCGAAGAGCGGCCGAC 382
Db	
Qy	294 CCAGAAATTCCTCTCTGAGGCAGACATGCAAGTAGACTCGATAGAGAGCTCCGGTC 353
Db	
Qy	393 AGAGCAGAAATACCTTCCACAGGAAGGAGACTGATACATGCAAAATTTCTTCAGGACAC 442
Db	
Qy	354 AAGTAGAATAACACCCCAAGGAAGGAGACTGATACATGCAAAATTTCTATCAGGACAC 413
Db	
Qy	443 ATGAAGGGGTGACTACTGGAAACGCAATTTCTTTATTGTTCCTCCAAACACAGATCAAAATAG 502
Db	
Qy	414 ATGATGGGATGACTACTGGGCACCAATTTCACTGCTTTTTCCTCCGACAGATCAAGAG 473
Db	
Qy	503 GCAGTGATCACCGGTGAATAAGCAATGTGTACCGACCTTCTCATGACAGACCAACTTATG 562
Db	
Qy	474 GCGGTGATTACAGTGAATGGCTAAGGCGTACAGACCATCCCATGCAGATGCAACCTATG 533
Db	
Qy	563 ACTTCAAGTACCGTGTAGAGCTGTACAGGAGGTGGAGGTCTCGGGCAGAAAAACCG 622
Db	
Qy	534 ACTTCAAGTATGTGTGTCAGAGCTGTGACGGAGGCGGAAGGTTCATCGGCCAGAGAAACCA 593
Db	
Qy	623 TTGGAAGGGGTGCTGACGGGGCCCTCCCAAGAAAAATTTCTTAAGCTCAAAATGTGGATTAG 682
Db	
Qy	594 TTGCGAGGGTGGCTGCGAGGAGCTTCTTGCAAAAGAAAAATTTCTGAAGCTCAATTCAGGATGG 653
Db	
Qy	683 AGATCTTGTGCTTGTGTTTCCAAAGTGCATCAGGTGTGTGCTCCCGAGAGACGCGGTTGATT 742
Db	
Qy	654 AGATCTTGGCATTTGTTTCTTAAAGTGCATCAAGTCGTACTTCCAGAGAGATGCAGTTCATT 713
Db	
Qy	743 ATGGGTCTGTAACTTTTGGAAACAGATAGAGACACATCGTTAGATGTCCTGATCCAGAT 802
Db	
Qy	714 ATGAGACTGTAGCTTTGGAAACAGATAGAGACCAATTTGTAGATGTCCTGATCCAGAT 773
Db	
Qy	803 ACCGAGAGAGATGATAGACGCAATCGACAGGTACGAGTT 843
Db	
Qy	774 ATGCAGAGAGATGATTGCTGCCATTGATAAGTACGAGTT 814
Db	
RESULT 9	
LOCUS	CNS0ACYT 1538 bp mRNA linear HTC 06-FEB-2004
DEFINITION	Arabidopsis thaliana Full-length cDNA Complete sequence from clone GS102H06 of Adult vegetative tissue of strain col-0 of Arabidopsis thaliana (thale cress).
ACCESSION	BX814847
VERSION	GI:42472838
KEYWORDS	HTC; GS1T cDNA.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
REFERENCE	Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M. Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 1538)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT	The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. UNGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full length http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=arabidopsis. Location/Qualifiers 1..1538 /organism="Arabidopsis thaliana"
FEATURES	
source	

```
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTSL102H06"
/tissue_type="adult vegetative tissue"
/plasmid="pCWSF06"
complement(1..1538)
/gene="At1g48850"

ORIGIN
Query Match      30.7%; Score 499.8; DB 3; Length 1538;
Best Local Similarity 66.6%; Pred. No. 6.7e-134;
Matches 759; Conservative 0; Mismatches 377; Indels 3; Gaps 3;

QY 216 CCTAGAGGTGAAGGACATCGAAACACAGTGTGGAACACTCTTTCAGGTTGCAACCTATGG 275
DB 269 CTTCCAGATACAGCTACTGGAAGTTCATATGGACTCATTTTCGAGTTTCAACTTTGG 328

QY 276 TGAATCTCATGGGGTGGTGTGTTGTTGTTTATCATGTTGTTCCACCTAGAAATCCACT 335
DB 329 AGAATCACATGAGGAGGAGTTGTTGTTATCATTTGATGTTGTTCTCTCGTATTCCTACT 388

QY 336 CACTGAGGACAGCTACAAGTTGAATCGATCGAAGACGGCCCGGACAGAGCAGAAATAC 395
DB 389 TACTGAATCTGATTTGCAATTCGATCTCGATAGAGGAGGCGCTGTGTGAGCAGGATCAC 448

QY 396 CTCACAAGGAAGGAGCTATACATGCAAAATTCGTCAGGACACATGAAGGGGTGAC 455
DB 449 AACTCTAGAAAGAGACTGATCTGCGGATATCGTCTGAGTCACTGAAGGAATGAC 508

QY 456 TACTGGAAGCCCAATCTTGTGTTATGTTCCCAACACAGATCAATAGGACGATGATCACCG 515
DB 509 GACAGGAACACCTATCCATGTTGTTGTTACCAACACAGATCAGAGGAGTGTGATTACAG 568

QY 516 TGAATAGCAATGTGTACCGACCTTCTCATGAGACGCAACTTATGACTCAAGTACGG 575
DB 569 TGAATGTGCGTTGCTATAGACCATCGATGCTGATGCAACTTATGATCAATGAATATGG 628

QY 576 TGTAGAGCTGTACAGGAGGTGGAGGTCTCGGCGCAAAACCGTTGGAAGGTGCG 635
DB 629 TGTAGATCAGTGAGGTGGAGGAGATCTTACGTAGAGACCATTTGGGAGAGTTGC 688

QY 636 TGAGGGGCGCTCCCAAGAAATTTTAAGCTCAATGTTGATTTAGAGATCTTCTGCTT 695
DB 689 TCTGAGCTTTGGCCAG- AATTTTGAAGCAATTTGAGGACATGATGATTCGCTCTA 747

QY 696 TGTTTCCAAAGTGATCA- GGTGTGCTCCCAAGAGACGCGTTGATTTATGGTCTGTAA 754
DB 748 TGTCTCGCAAGTCCACCATTTGTTACTCCCAAGAAATTTGTAGACACGAGAAATTTA 807

QY 755 CTTTGGACAGATAGAGACATCGTTAGATGTTCTGATCCAGATACGACGAGAGA 814
DB 808 CACTCGAACAGATAGAAATAACATTTGTCAGATGCTTAATCCCGAGTATGCGAAAGA 867

QY 815 TGATAGACGCAATCGACAGATACAGTTTCAGGGGATTCGCTCGGTGAGTGATCAAT 874
DB 868 TGATAGCTCGGATGATGCTGTGACAGCAAAAGGAACTCTGTTGGTGTGTTGACT 927

QY 875 GGTTCGCTAGAAACGTTCTCGCGGGCTGGTTCCTGTTGTTGTTGTTGTTGTTGTTGTT 934
DB 928 GCAATGTTTCGGAATGCTCCACGTGGCTTGGTACACCGGTTTTCGATAAATTTGAAGCAG 987

QY 935 AACTGGCAAAAGCTATGTTCTTATCTCTGCGAGCAACGGGTTTCAGATTCGACCGAT 994
DB 988 AACTGGCAAAAGCTTGTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1047

QY 995 TCTCGGACCGACTTTGACAGGAAGTGAGCATAATGATGAGTTTATATGATGATGAGGCTG 1054
DB 1048 TTGACAGTACCTTTTGTACTGTTGATGATGATGATGATGATGATGATGATGATGATG 1107

QY 1055 GAAAGTTCAGGACACGGAATAATCGCTCGGGTGGTGTGTCAGGAGGGGATATGGAATGTTG 1114
DB 1108 GAAAGTACGTACCAAGAACCAACCGATCTGGTGGAAATTCAGGGAGGGATCTCAATGTTG 1167
```

```
1115 AGATTGTGCACTTCAAGGTGCTTTTAAGCCGAC-ACCATCTATCGGGTGAACAGAAC 1173
DB 1158 AATAATAAATAGAGTAGCTTCAAGCCAAATCAATCAATTTGGAAGGAGAAAT 1227
QY 1174 ACCGTGTCAAGGGAGCGTCAAGACGTTTGAGCTTCTAGCAAGAGGCGCCATGACCATGC 1233
DB 1228 ACGGTAAACAGAGACAAGGTAGAAACCGAAATGATTGCGCGTGGTCTCATGATCCTTGT 1287
QY 1234 GTCGCCCCCTCAGAGCTGTTCTGTGTGGAATCCATGCGCGGTGTTGTTCTCATGACACG 1293
DB 1288 GTTGTTCACGAGCTGTGCAATGTTGGAAGCAATGTTGGTCTAGTTCTTGTGATCAA 1347
QY 1294 CTGATGGCGCAGCTGGCTCAGTGGCAGATGTTGCGCTCAATGCTGCACTTCAAGAAC 1352
DB 1348 TTGATGGCGCAATACGCAATGCCATTTGTTTCCATAAATCCAGAGTTGCGGAACC 1406

RESULT 10
CF303160 723 bp mRNA linear EST 15-AUG-2003
LOCUS ABF1--01-104.g1 ABF3-overexpressing transgenic rice lambda phage
DEFINITION cDNA library (ABF1) Oryza sativa (japonica cultivar-group) cDNA
clone ABF1--01-104, mRNA sequence.
ACCESSION CF303160
VERSION CF303160.1 GI:33674921
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 723)
AUTHORS Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C.,
Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.
TITLE Large-scale Sequencing Analysis of Rice ESTs
JOURNAL Unpublished (2003)
COMMENT Contact: Nahm B.H.
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
of Bioeconomics and Bioinformatics, Myongji University
Yongin, Gyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bnhnm@gbio.com, bnhnm@bio.myongji.ac.kr.

FEATURES
Location/Qualifiers
1..723
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:39947"
/clone="ABF1--01-104"
/tissue_type="leaf"
/dev_stage="14 days after germination"
/lab_host="E.coli SOLR"
/clone_lib="ABF3-overexpressing transgenic rice lambda
phage cDNA library (ABF1)"
/notes="vector: pbluescript SK(+); Site1: EcoRI; Site 2:
lambda Uni-ZAP XR vector at 5' end with EcoRI and 3' end
with XhoI site. mRNA was prepared from ABA-responsive
element binding transcription factor 3 overexpression
line."

ORIGIN
Query Match      30.7%; Score 498.6; DB 7; Length 723;
Best Local Similarity 83.5%; Pred. No. 1.2e-133;
Matches 589; Conservative 0; Mismatches 114; Indels 2; Gaps 2;

QY 343 GCAGACCTACAGTTGAACTCGATCGAAGACGGCCCGGACAGAGCAGAAATTAACCTCCACA 402
DB 20 GCAGATATGCAAGTAGAACTCGACCGGAGACGGCCGACGAGCAGAAATTAACCCCA 79
QY 403 AGGAAGGAGACTGATACATGCAAAATTTCTGTGAGGGACACATGAAGGGGTGACTACTGGA 462
```


Db	80	AGAAAGGAGACTGACACTTGCANAAATCTTTTCAGGGACACATGAAGGAATGACCACCTGGG	139
Qy	463	ACGCCAAATCTTGTATTGTCTCCCAAAACACAGATCAAAATAGGCAGTGTACCGTGAATA	522
Db	140	ACACCAATTCATGTTTTGTCTCCGGAACACAGATCAGAGAGGGGTGATTACAGTGAATG	199
Qy	523	GCCAAATGTGTACCGACCTTCTCATGCAGACGCCAATTTATGATCTCAAGTACGGTGTAGA	582
Db	200	GCTAAGGCCTACAGACCTTCACTGTGCAGATCAACTTATGACTTTCAATACGGTGTAGA	259
Qy	583	GCTGTACAGGAGGTGGGAGTCTCTCGGGCAGAAAAACCGTTGGAAGGCTGGCTGCAGGG	642
Db	260	GCAGTCAGGGAGGTGGAGATCATCAGCAAGAGAGACCAATTGNAAGGTGGCTGCAGGA	319
Qy	643	GCCCTCCCAAGAAAAATCTTTAAGCTCAATGTGGATTAGAGATCTTGTCTGTTGTTCCT	702
Db	320	GCTCTTGCNAAAGAAAAATCTTTAAGCTCAAAATCTGGAGTAGAGATCTTTGGCGTGTGTGCC	379
Qy	703	AAAGTGATCAGGTTGTGTCTCCCAAGAGACGCGGTGATTATGGTCTGTAACTTTGGAA	762
Db	380	AAGGTGCATCAAGTTGTACTACCAAGAGATGCCGTTGATTATGACACTGTAAACAATGGA	439
Qy	763	CAGATAGAGACAACATCGTTAGATGTCTGATCCAGATGACGAGACGAGAAGATGATGAC	822
Db	440	CAGATAGAAGCAACATTTGTTAGATGTCTCTGATCCAGAATATGCACAGAAGATGATGAT	499
Qy	823	GCAATCCACACAGATACAGATTTCGAGGGGATTCGGTTCGGTG - GAGTGATACATCGCTCGC	881
Db	500	GCAATCGATAAAGTACGAGTTTAGAGGTGATTCGATTTGGTGGGTGTGTGCATCATGCAATGC	559
Qy	882	TAGAAAAGTTCCCTCGCGGGCTCGGTTCTCTGTGTTTCGACAAGCTCGAATCCGAACCTGGC	941
Db	560	AAGAAATGTTCTCTCGTGGGATTGGCGTCTCTGTATTTGACAAAATTGAGGCTGAATGGC	619
Qy	942	AAAAGCTATGCTTTCTATTCTCGCGAGCAACGGGTTCGAGATTGCGACGGATTCGCGG	1001
Db	620	GAAAGCTATGCTTTCTCTTCTCTGC - AGCAAGGGTTTTGAGATCGCGAGTGTGATTCGAGG	678
Qy	1002	GACCGACTTGACAGGAAGTAGCAATAATGATGAGTTTTTATATGA	1046
Db	679	TACTGACTACACTGGAAGTAGCAATAATGATGAGTTCTATATGA	723

RESULT 11	
BJ268020	
LOCUS	
DEFINITION	BJ268020 Y. Ogiwara unpublished cDNA library, Wh_oh Triticum aestivum CDNA clone whonl4j05 5', mRNA sequence.
ACCESSION	BJ268020
VERSION	BJ268020.1 GI:20083164
KEYWORDS	EST.
SOURCE	Triticum aestivum (bread wheat)
ORGANISM	Triticum aestivum
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.

REFERENCE	1. (bases 1 to 710)
AUTHORS	Ogihara,Y. and Murai,K.
TITLE	Expressed genes in <i>Triticum aestivum</i>
JOURNAL	Unpublished (2002)
COMMENT	Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp.
FEATURES	Location/Qualifiers
source	1..710 /organism="Triticum aestivum" /mol_type="mRNA" /cultivar="Chinese Spring"
Db	661 CATTGGTCTCATGGACCAGCTGATGGACATGTTGCTCAGTGCCGAGATG 710
RESULT 12	
LOCUS	BU099249 702 bp mRNA linear EST 29-AUG-2000
DEFINITION	WHE3305_C09_E17ZS Chinese Spring wheat drought stressed root cDNA library_Triticum aestivum CDNA clone WHE3305_C09_E17, mRNA sequence.
ACCESSION	BU099249
VERSION	BU099249.1 GI:22546938
KEYWORDS	EST.
SOURCE	Triticum aestivum (bread wheat)
ORGANISM	Triticum aestivum
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Poideae; Triticeae; Triticum.
 1 (bases 1 to 702)
 REFERENCE
 AUTHORS Anderson, O.D., Chao, S., Close, T.J., Crossman, C., Fenton, R.D., Lazo, G.R., Nguyen H.T., Pham, J., Rausch, C.J., Turuspekov, Y., Wilson, C., Woo, J., and Zhang, D.
 TITLE The structure and function of the expressed portion of the wheat genomes - Chinese Spring drought stressed root cDNA library
 JOURNAL Unpublished (2002)
 COMMENT Contact: Olin Anderson
 US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
 800 Buchanan Street, Albany, CA 94710, USA
 Tel: 5105959773
 Fax: 5105959818
 Email: oanderson@pw.usda.gov
 Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
 Seq primer: SK primer.

FEATURES

Location/Qualifiers
 1..702
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="Chinese Spring"
 /db_xref="taxon:4565"
 /clone="WHE3305 C09 E17"
 /tissue_type="root"
 /dev_stage="Full tillering stage"
 /lab_host="E. coli SOUR"
 /clone_lib="Chinese Spring wheat drought stressed root cDNA library"
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site 1: EcoRI; Site 2: XhoI; Plants were grown under normal conditions, then drought stressed to 80%, 70% and 60% RWC at Texas Tech University (D. Zhang in H.T. Nguyen lab). Total RNA was prepared separately for roots collected at the three different drought conditions. Equal amounts of total RNA were pooled from all three samples, poly(A) RNA were purified, one cDNA library was made, and the cDNA clones were in vivo excised to give pBluescript SK(-) phagemids in the TJ Close lab at the University of California, Riverside (Fenton, Turuspekov). Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."
 ORIGIN
 Query Match 29.5%; Score 480.4; DB 5; Length 702;
 Best Local Similarity 80.5%; Pred. No. 2.5e-128;
 Matches 562; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 634 GCTGAGGGGCGCTCCCAAGAAAATCTTAAGCTCAAAATGGATTAGAGATCTTGTGC 693
 DB 5 GCTGAGTAGCTGTGGCAAGAAAATCTTAAGCTCAAAATGGATTAGAGATCTTGTGC 64
 QY 694 TTGTGTTCCAAAGTGCATCAGGTGTGCTCCAGAGACGCGGTGATTATGGGTCGTGA 753
 DB 65 TTGTGTTCCAAAGTGCATCAGGTGTGCTCCAGAGACGCGGTGATTATGGGTCGT 124
 QY 754 ACTTTGGAACATAGAGAGCAACATCGTTAGATCTCTGATCCAGAGTACCCAGAGAG 813
 DB 125 ACCCTGGATCAGATAGAGAGCAACATTTGTAGATCTCTGATCCAGAGTATGCAGAAG 184
 QY 814 ATGATAGAGCAATGCAGAGTACGAGTTCGAGGGGATTCGGTGGTGGATGATCACA 873
 DB 185 ATGATGATGCAATGATAAAGTACAGATTATGGGAATTCGATTTGGTGGGTGTGTACA 244
 QY 874 TCGTCGCTAGAAAGCTTCTCCGCGGCTCGGTCTCTCTGTGTTCCGACAAAGCTCGAATCC 933
 DB 245 TGCATTGCGAAGATGTTCTCTGCGGCTTGACTCTCTGTATTATGACAACTGAAGCT 304
 QY 934 GNACTGGCAAAAGCTATGCTTTCTATTCTTCGAGCAACAGGGTTTCGAGATTGGCAGCGGA 993
 DB 305 CTACTGCGCAAGGCTATGCTTTCTCTCTCTGCAAGCAAGGGGTTTGAGATCGGTAGTGA 364

QY 994 TTGCGCGGACCGACTTCACAGGAAGTGCAGATAATGATGAGTTTTATATGATAGGCT 1053
 DB 365 TTTGAGGAGTACTGACCTTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 424
 QY 1054 GGAAGTGTGACGACACCGACTAATCGTCGGGTGTGTCAGGGAGGAGATATCGAATGTT 1113
 DB 425 GGAAGTGTGACGACACCGACTAATCGTCGGGTGTGTCAGGGAGGAGATATCGAATGTT 484
 QY 1114 GGAAGTGTGACCTTCAAAAGTGTGTTTAAAGCCGACACCATCTATCGGGGTGAACAGAC 1173
 DB 485 GAAACTATATATCTTCAAAAGTGTGTTTAAAGCCGACACCATCTATTTGGGAAGAGCAAA 544
 QY 1174 ACCGTGTCAAGGGAGCGTCAGAACCTTCAGCTTCTAGCAAGAGGCGGCATGACCATGC 1233
 DB 545 ACTGTAAACAGGGATCATGAGGATATCGAATCTTCTGACAAAGGGTGCATGACCATGTT 604
 QY 1234 GTGCGCCCTCGAGCTGTTCTGTGTGGAATCCATGCGCGGTGGTCTCTCATGACACCA 1293
 DB 605 GTTGTCCCTCGGCTGTTTCCAATGTTGAGAGACGATGCTGCTCTCATGACACCA 664
 QY 1294 CTGATGGCGCACCTGGCTCAGTGGAGAGATGTTCCGCT 1331
 DB 665 CTGATGGCACATGTTGCTCAGTGGAGAGATGTTCCGCT 702

RESULT 13

CAL89902
 LOCUS SCCLRLIC06B07.g LR1 Saccharum officinarum cDNA clone SCCLRLIC06B07
 DEFINITION 5', mRNA sequence.
 ACCESSION CAL89902
 VERSION CAL89902.1 GI:35133533
 KEYWORDS EST.
 SOURCE Saccharum officinarum
 ORGANISM Saccharum officinarum
 REFERENCE 1 (bases 1 to 864)
 AUTHORS Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
 TITLE The libraries that made SUCEST
 JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
 COMMENT Contact: Arruda P
 Centro de Biologia Molecular e Engenharia Genetica
 Universidade Estadual de Campinas
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil
 Tel: 55 19 3788 1137
 Fax: 55 19 3788 1089
 Email: parruda@unicamp.br
 Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at http://www.bcccenter.fcav.unesp.br
 Plate: C06 row: B column: 07
 Seq primer: T7 Promoter Primer.
 Location/Qualifiers
 1..864
 /organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"
 /clone="SCCLRLIC06B07"
 /lab_host="DH10B"
 /clone_lib="LR1"

FEATURES

source
 /note="Organ: Leaf roll from field grown adult plants (large insert library); Vector: pSport1; Site 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [leaf roll from field grown adult plants (large insert library)]. cDNA was prepared from polyA+ mRNA using Superscript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be

ORIGIN obtained at <http://eucest.lad.ic.unicaamp.br/public>

Query Match 29.4%; Score 477.4; DB 6; Length 864;
Best Local Similarity 82.5%; Pred. No. 2e-127;
Matches 547; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

Qy 701 CCAAGTGCATCAGGTTGCTCCAGAGAGCGGTTGATTGATGGTCTGAACCTTGG 760
Db 1 CTAAGTGCATCAAGTCTACTTCCAGAGATGCAAGTTGATTGAGACTGTAACTTGG 60
Qy 761 AACGATAGAGAGCAACATCGTTAGATGCTCCTGATCCAGAGTACGAGAGAGATGATAG 820
Db 61 AACGATAGAGAGCAACATCGTTAGATGCTCCTGATCCAGAGTATGAGAGAGATGATG 120
Qy 821 AGGCAATCGACAGAGTACAGTTGAGGGGATTCGGTGGTGGAGTATCATCGCGTCG 880
Db 121 CTGCCATTGATAAGTACAGTTGAGGAGATTCAAATTTGGTGGGTCTGCATGCAATTG 180
Qy 881 CTAGAAAGTTCTCCGCGGCTCGGTTCTCTGTTTCGACAAGCTCGAATCCGAATCGG 940
Db 181 CAAGAAAGCTCCCTCGTGGTCTTGGCTCTCTGTTTGGACAACTTGAAGCTGAACCTGG 240
Qy 941 CAAAGCTATGCTTTCTATTCCTCGAGCAACGGGTTCCAGATTGGCAGCGGATTCGCG 1000
Db 241 CAAAGCTATGCTTTCTCTTCTGCAAGAGGGGTTGAGATTGGCAGTGGTTGGTG 300
Qy 1001 GGACCGACTTGACAGAGAGTGAAGCAATATGATGATGTTTATATGATGAAGCTGGAAAGTG 1060
Db 301 GTACAGACTTTTACTGGAAGTGAAGCAACGATGAGTTCTATATGATGAGGCTGGAAATG 360
Qy 1061 TCAGACACGAGCTAATCGCTCGGTTGTTGTCAGGGAGGATATCGAATGTTGAGATTG 1120
Db 361 TCGCAGACAGCACTAATCGCTCAGCGGTTGTCAGGGAGGATATCGAATGTTGAAATTA 420
Qy 1121 TGCCTTCAAGTTGCTTTTAAAGCCGACCACTTATCGGGGTGAAAAGAACACCCGTTG 1180
Db 421 TTTACTTCAAGTGGCTTTTAAAGCCACAGCAACTATCGGAAGAAACAAATACTGTGT 480
Qy 1181 CAAAGGAGCTCAGAACGTTGAGCTTCTAGCAAGAGGGCGGATGACCCATCGCTCGGCC 1240
Db 481 CAAAGGAGCATGAGATGTTGAACCTTTTGGCAAGGGGCGCATGACCCGCTGTGTGTC 540
Qy 1241 CTCGAGCTGTTCTCTGTGTGGAATCCATGCGCGCTGCTCTCATGACCAAGCTGATGG 1300
Db 541 CTCGAGCTGTTCTATGTTGGAATCAATGCTGCTGCTGCTGCTGATGACCAAGCTCATGG 600
Qy 1301 CGCAGTGGCTCAGTGCAGATGTTTCGGCTCAATPACTGCACTTCAAGAACCAAGTTGGCT 1360
Db 601 CGCATATTGGTCAATGTGAGATGTTTCGGGTGAACCTTGGCCCTACAGAGCCCATGTT 660
Qy 1361 CTT 1363
Db 661 CTT 663

RESULT 14
CD870292
LOCUS AZ02.113P12F001120 AZ02 Triticum aestivum cDNA clone AZ02113P12,
DEFINITION mRNA sequence.
ACCESSION CD870292.1 GI:32554108
VERSION CD870292
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 694)
AUTHORS Genoplante.
TITLE Genoplante, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)

COMMENT Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (<http://www.genoplante.com>)
and <http://genoplante-info.infobiogen.fr>.

FEATURES
source
Location/Qualifiers
1..694
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="AZ02113P12"
/tissue_type="root"
/clone_lib="AZ02"

ORIGIN
Query Match 29.2%; Score 475.6; DB 6; Length 694;
Best Local Similarity 81.3%; Pred. No. 6.2e-127;
Matches 564; Conservative 0; Mismatches 129; Indels 1; Gaps 1;

Qy 331 CCACCTCACTGAGGAGACCTTACAAAGTTGAACTCGATCGAAGAGCGGCCGCGACAGCAGA 390
Db 2 CCACCTCACTGAGGAGACATGCAAGGAGACCTTGTATCGAAGGCGGCCAGGTCAAGCAGA 61
Qy 391 ATAACTCTCCAAGAGAGGAGACTGATACATATGCAAAATTTCTGTGAGGAGACATATGAAGG 450
Db 62 ATAACAACCCCAAGAGGAGAGCGGATCTTGTAAATTTCTTTTCAAGGAGCATATGAAGGA 121
Qy 451 GTGACTACTGGAAGCGCAATTTCTGTTTATGTCCTCCAAACACAGATCAAAATGAGCAGTAT 510
Db 122 ATGACCACTGGGAGCGGATTCATGTTTGTCTCCAAACACGATCAAGAGGGGGTAT 181
Qy 511 CACCGTGAATAGCAATGTGTACCGACCTTCTCATGACAGACGCAACTTATGACTTCAAG 570
Db 182 TACACTGAATAGCTAAGGCGTACAGACCTTCCCATGCGGATTTGACTTATGACCTCAAG 241
Qy 571 TACGTTGTTAGAGCTGTACAGGAGGTGGAGGTCTCTCGGGCAGAGAAAACCGTTGGAGG 630
Db 242 TACGTTGTTAGATCTGTTCAGGAGGTGGAAGGTCTATCGGCAAGAGAAAACCATTTGGAAG 301
Qy 631 GTGCTCAGGGGCGCTCCCAAGAAAATTTCTTAAGCTCAAAATGTGATTAGAGATCTTG 690
Db 302 GTAGCTCAGAGCTGTTCGAAGAAAATTTCTTAAGCTGAATGTGAGTAGAGATTCTA 361
Qy 691 TCGTTTGTTCCTCAAGTGCATCAGGTTGTCTCCCAAGAACGCGGTTGATTATGGGTCT 750
Db 362 GCATTTGTTTCCAAAGTGCATCAAGTGTACTTCTCTGAAGACGCGAGTTGATTGAAACT 421
Qy 751 GTAACTTTGGAAACAGATAGAGAGCAACATCGTTAGATGTCTGTATCCAGAGTACGACAG 810
Db 422 CTTACCTGGATCAGATAGAGAGCAACATTTGTAGATGTCTCTGATCCAGAAATATGCACAG 481
Qy 811 AAGATGATAGACCAATCCGACAGATACAGATTTCGAGGGGATTCGGTGGAGTGCATC 870
Db 482 AAGATGATGATCAATTTGATAAAGTACGAGTTAATGGGAATTCGATTGGTGGGGTGGTC 541
Qy 871 ACATGCTGCTGAAGAACGTTCTCTCGGGGCTCGGTTCTCTCTGTTCGACAAAGCTCGAA 930
Db 542 ACATGCTATGCCAGAAATGTTCTCTCGTGGGTTGGCTCTCTCTGTATTTGACAACTTGAA 601
Qy 931 TCCGAACCTGGCAAGGCTATGCTTTTATTCCTGCGAGAGCAACGGGTTTCGAGATGGGAGC 990
Db 602 GCTCTACTGGCAAGGCTATGCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
Qy 991 GGATTCGCGGAGCCGCTTTGACAGGAAGTGAGC 1024
Db 661 GGATTCGAGTACTGACCTAACTGGGAAGTGAGC 694

RESULT 15

CA124299 690 bp mRNA linear EST 23-SBP-2003
 LOCUS SCQGLR1086A09.g Lr1 Saccharum officinarum cDNA clone SCQGLR1086A09
 DEFINITION 5' mRNA sequence.
 ACCESSION CA124299
 VERSION 1 GI:34977607
 KEYWORDS EST.
 SOURCE Saccharum officinarum
 ORGANISM Saccharum officinarum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
 complex.
 1 (bases 1 to 690)
 Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
 The libraries that made SUCEST
 Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
 Contact: Arruda P
 Centro de Biologia Molecular e Engenharia Genetica
 Universidade Estadual de Campinas
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil
 Tel: 55 19 3788 1137
 Fax: 55 19 3788 1089
 Email: parruda@unicamp.br
 Clone distribution: clone distribution information can be found
 through the Brazilian Clone Collection Center (BCCC) at
 http://www.bcccenter.fcav.unesp.br
 Plate: 086 row: A column: 09
 Seq primer: 17 Promoter Primer.
 Location/Qualifiers
 1..690
 /organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"
 /clone="SCQGLR1086A09"
 /lab_host="DH10B"
 /clone_lib="Lr1"
 /note="Organ: Leaf roll from field grown adult plants
 (large insert library); Vector: pSport1; Site 1: SalI;
 Site 2: NotI; An unidirectional cDNA library generated
 from [leaf roll from field grown adult plants (large
 insert library)]. cDNA was prepared from polyA+ mRNA
 using SuperScript Plasmid System Kit (Invitrogen). The
 double-strand cDNAs were fractionated in a sepharose
 CL-2B 40cm-columns and fragments sizing between 0.8 and
 1.5 Kb were directionally cloned into the vector. Details
 of each source of RNA and library construction can be
 obtained at http://sucest.lad.ic.unicamp.br/public"

FEATURES

source
 1..690
 Search completed: August 27, 2005, 04:20:17
 Job time : 4912 secs

ORIGIN

Query Match 29.1%; Score 473; DB 6; Length 690;
 Best Local Similarity 82.2%; Pred. No. 3.6e-126;
 Matches 567; Conservative 0; Mismatches 121; Indels 2; Gaps 2;
 QY 538 CTTCTCATGACAGCGCAACTTATGACTTCAAGTACGGTGTAGAGCTGTACAGGAGGT 597
 DB 1 CCATCCCATGACAGTCAACCTATGACTTCAAGTATGGTGTGACAGCTGTGACGGAGGC 60
 QY 598 GGGAGGTCTCGGGCAGAGAAAACCGTTGGAAGGGTGGCTGCAAGGGGCCCTCCCAAGAAA 657
 DB 61 GGAAGGTCTATCGCCAGAGAAACCAATTGGCAGGGTGGCTGCAAGGAGCTCTTGCAAGAAA 120
 QY 658 ATTCTTAAGCTCAATGTGGATTAGAGATCTGTGTTGTTTCCAAAGTGCATCAGGTT 717
 DB 121 ATTCTGAAGCTCAATTCAGGAGTGAGATTTTGGCATTTGTTCTAAAGTGCATCAAGTC 180
 QY 718 GTGCTCCAGAGAGACCGGTTTGATTATGGTCTGTAACTTTTGGAAACAGATAGAGCAAC 777
 DB 181 GTACTTCCAGAGAGTGCAGTTGATTATGAGACTGTAACTTGGAAACAGATAGAGCAAC 240
 QY 778 ATCGTTAGATGTCCTGATCCAGAGTACCGAGAGAGATGATAGACGCAATCGACAGATA 837
 DB 241 ATTGTTAGATGTCCTGATCCAGAAATATGCAGAGAGATGATTGCTGCCATTGATAAGTA 300

QY 838 CGAGTTTCAGAGGGATTTCGGTGGTGGAGTGTATCATCGTGCCTAGAAAAGTTTCTCGC 897
 DB 301 COAGTTAGAGAGATTTCANTTGGTGGGTGTCATCATGATGCAAGAAACGTCCCTCGT 360
 QY 898 GGGCTCGGTTTCTCCTGTGTTTCGACAAAGCTCGAATCCGAACCTGGCAAAAGCTATGTTTCT 957
 DB 361 GGTCTTGGCTCTCCTGTGTTTTCGACAAACTTGAAGCTGAACCTGGCAAAAGCTATGTTTCT 420
 QY 958 ATTCTCGAGCAACGGGTTTCGAGATTGGCAGCGGATTCGCCGGGACCGACTTGAAGATA 1017
 DB 421 CTTCTCGCAAGCAAGGGGTTTGAGATTGGCAGTGGGTTTGTGGTACGAGCTTTACTGGA 480
 QY 1018 AGTGAGCATATGATGAGTGTATATATGATAGGCTGGAAGTGTCCAGCACCGGACTAAT 1077
 DB 481 AGTGAGCATATGAGTGTATATATGATAGGCTGGAAGTGTCCAGCACCGGACTAAT 540
 QY 1078 CGCTCGGTTGCTGTCAGGGAGGGATATCGAATGTGTAGATTGTGCATTTCAAGTTGC- 1136
 DB 541 TCCTCAGCGGTGTTTCAGGGAGGGATATCGAATGTGTAGATTGTGCATTTCAAGTTGC 600
 QY 1137 TTTTAAGCCGACACCATCTATCGGG-GTGAACACAGAACACCGTGTCAAGGGAGCGTCA 1195
 DB 601 TTTTAAGCCAAACAGCAACTATCGGGAGGAAACAAAATACTGTGTCAAGGGAGCATGANG 660
 QY 1196 ACGTTGAGCTTCTAGCAAGAGCGGCCCATG 1225
 DB 661 ATGTGGAATTTTGGCAAGGGGGGCCCATG 690

THIS PAGE BLANK (USPTO)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1626	100.0	1626	4	US-09-743-207-7	Sequence 7, Appli
2	870.4	53.5	1635	4	US-09-743-207-1	Sequence 1, Appli
3	621	38.2	366	4	US-09-743-207-11	Sequence 11, Appli
4	500.6	30.8	1015	4	US-09-743-207-5	Sequence 5, Appli
C 5	274	16.9	1053	3	US-09-610-040-1	Sequence 1, Appli
6	274	16.9	1053	3	US-09-610-040-7	Sequence 7, Appli
C 7	274	16.9	1053	4	US-10-267-763-1	Sequence 1, Appli
8	274	16.9	1053	4	US-10-267-763-7	Sequence 7, Appli
9	203.8	12.5	1080	4	US-09-248-796A-4099	Sequence 4099, Ap
10	185	11.4	300	4	US-09-313-294A-4734	Sequence 4734, Ap
C 11	180.2	11.1	640681	4	US-09-790-988-1	Sequence 1, Appli
12	179.6	11.0	1089	4	US-09-543-681A-120	Sequence 120, App
13	178	10.9	1664976	4	US-08-916-421B-1	Sequence 1, Appli
14	178	10.9	1664976	4	US-09-692-570-1	Sequence 1, Appli
15	170	10.5	1239	3	US-09-064-693A-20	Sequence 20, Appl
16	170	10.5	4530	3	US-09-064-693A-26	Sequence 26, Appl
C 17	169.6	10.4	6464	3	US-09-221-017B-168	Sequence 168, Appl
18	169.4	10.4	541	4	US-09-743-207-13	Sequence 13, Appli
19	165	10.1	1104	4	US-09-252-352-3536	Sequence 3536, Ap
20	160	9.8	1233	4	US-09-252-991A-13477	Sequence 13477, A
C 21	160	9.8	1263	4	US-09-252-991A-13857	Sequence 13857, A
22	159.6	9.8	1134	4	US-09-489-039A-5375	Sequence 5375, Ap
23	156.4	9.7	1140	4	US-09-252-991A-13588	Sequence 13588, A
24	156.4	9.6	1830121	4	US-09-557-884-1	Sequence 1, Appli
25	156.4	9.6	1830121	4	US-09-643-990A-1	Sequence 1, Appli
26	139.8	8.6	479	4	US-09-743-207-9	Sequence 9, Appli
27	136.2	8.4	21410	4	US-09-596-002-12	Sequence 12, Appl

Qy 915 GTTCGACAAAGCTCGAATCCGAACCTGCAAAAGCTATGCTTTCTATTCCTCGAGCAACGG 974
Db 875 ATTTGACAAACTTGAGGCTGAATTGGCGAAGCTAATGCTTTCTCTTCTCGAAGCAACGG 934
Qy 975 GTTCGAGATTGGCAGCGGATTGG 997
Db 935 GTTTCGAGATCGGCACTGGGATTG 957

RESULT 4

US-09-743-207-5
; Sequence 5, Application US/09743207
; Patent No. 6653531
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Pember, Stephen O.
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-A
; CURRENT APPLICATION NUMBER: US/09/743,207
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: PCT/US99/16353
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1015
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-743-207-5

Query Match 30.8%; Score 500.6; DB 4; Length 1015;
Best Local Similarity 80.3%; Pred. No. 1.3e-151;
Matches 587; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

Qy 631 GTGGCTGCAGGGCCCTCCCAAGAAATCTTAAGCTCAATGTGGATTAGAGATCTTG 690
Db 5 GAGGCTGCAGGAGCTGTGTCAAGAAATCTTAAGCTGAAATGTGGAGTAGAGATTTCTA 64
Qy 691 TCGTTTGTTCAAAGTCATCAGGTGTGCTCCCAAGAGACGCGTTTGATTATGGGTCT 750
Db 65 GCATTTGTTCAAAGTCATCAAGTGTGCTCTTGAAGCGCAGTTGATTATGAACT 124
Qy 751 GTAATTTGGAAACAGATGAGAGCAACATCGTTAGATGCTCTGATCCAGAGTACGCAGAG 810
Db 125 CTACCTCGATCAGATAGAGAGCAACATTTGTAGATGCTCTGATCCAGAAATATGCACAG 184
Qy 811 AAGATGATAGACCAATCGACAGAGTACGAGTTCAGGGGATTCGGTCGGTGGATGATC 870
Db 185 AAGATGATGATGCAATTTGATTAAGTACGAGTTAATGGGAATTCGATTTGGGGTGGTC 244
Qy 871 ACATCGCTCGCTAGAAAGCTTCTCTCGGGGCTCGGTTCTCTGTGTTCCGACAAAGCTCGAA 930
Db 245 ACATGCAATTCGAGAAATGTTCTCTGGGCTTGCTCTCTGTATTTGACAAACTTGAA 304
Qy 931 TCCGAATCGGCAAAAGCTATGCTTTCTATTCCTCGAGCAACGGGTTTCGAGATTGGCAGC 990
Db 305 GCTCTACTGGCAAGGCTATGCTTTCTCTCTCGCAAGCAAGGGGTTTGAGATCGGTAGT 364
Qy 991 GGATTCGCGGACCGACTTGACAGAGTGGCATATGATGATGATTTTATATGGATAAG 1050
Db 365 GGATTTGCGAGGTACTGACCTAACTCGAAGTGGCATTAACGATGATTTCTATATGGACGAG 424
Qy 1051 GCTGAAAGTGTGAGACACGGAATTAATCGCTCGGGTGGTGTGCGAGGAGGATATCGAAT 1110
Db 425 GCTGAAATGTGAAGAACACGACCAATCGCTCGGGCGGTGTACAGGGAGGATATCAAT 484
Qy 1111 GTTGAGATTGCACTTCAAAGTGTGCTTTTAAGCCGACACCATCTATCGGGGTGAAACAG 1170
Db 485 GGTGAAACTATATATCTTCAAAGTAGCTTTTCAAGCCCAACAGCAACTATTTGGGAAGAAGCA 544

Qy 1171 AACACCGTGTCAAGGGAGCGTCAAGCGTTGAGCTTCTAGCAAGAGGCGCCCATGACCCA 1230
Db 545 AATACTGTAAACAGGGATCATGAGGATATCAAACTTCTGACAAAGGGTCCGCATGACCCA 604
Qy 1231 TGGCTCGCCCTCGAGCTGTTCTCTGTGGTGGAAATCCATGGCGGTTGTGCTCATGGAC 1290
Db 605 TGTGTGCTCCTCGGGCTGTTCCAAATGTTGGAGACGATGCTGCAATTTGTTCTCATGGAC 664
Qy 1291 CAGCTGATGGCGACGCTGGCTCAGTGGCGATGTTGGCGCTCAATGCTGCACTTCAAGAA 1350
Db 665 CAGCTGATGGCACATGTTGCTCAGTGGCGATGTTCCGCTGAACCTCGCCCTACAAGAA 724
Qy 1351 CCAGTTGGCTC 1361
Db 725 CCAATCGGCTC 735

RESULT 5

US-09-610-040-1/C
; Sequence 1, Application US/09610040
; Patent No. 6465217
; GENERAL INFORMATION:
; APPLICANT: Boyes, Douglas
; APPLICANT: Davis, Keith
; APPLICANT: Woessner, Jeffrey
; APPLICANT: Gorlach, Jorn
; APPLICANT: Hamilton, Carol
; APPLICANT: Hoffman, Neil
; APPLICANT: Kloti, Andreas
; APPLICANT: Zayed, Adel
; APPLICANT: Ascenzi, Robert
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHASE
; FILE REFERENCE: 9128.14
; CURRENT APPLICATION NUMBER: US/09/610,040
; CURRENT FILING DATE: 2000-07-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1022)..(1023)
; OTHER INFORMATION: "n" indicated any nucleotide (A, C, T, or G).
; NAME/KEY: misc feature
; LOCATION: (1025)..(1025)
; OTHER INFORMATION: "n" indicated any nucleotide (A, C, T, or G).
; NAME/KEY: misc feature
; LOCATION: (1033)..(1034)
; OTHER INFORMATION: "n" indicated any nucleotide (A, C, T, or G).
; NAME/KEY: misc feature
; LOCATION: (1044)..(1044)
; OTHER INFORMATION: "n" indicated any nucleotide (A, C, T, or G).
US-09-610-040-1

Query Match 16.9%; Score 274; DB 3; Length 1053;
Best Local Similarity 64.0%; Pred. No. 6.1e-78;
Matches 430; Conservative 0; Mismatches 240; Indels 2; Gaps 1;

Qy 216 CCTAGAGGTGAGGCATCTCGAACAACAGTTTGGAACTACTTTTCAGGTTGCAACTATGG 275
Db 748 CTTCCAGATCAAGCTACTGGAAGTTCTATATGGGACTCATTTTCGAGTTTCAACTTTGG 689
Qy 276 TGAATCTCATGGGGTGGTGTGGTTGTGTTATCAGTGGTTGTCCACTAGAAATCCACT 335
Db 688 AGAATCATATGGAGGAGGTGGTTGTATCATTTGATGGTTGCTCTCTGTTATTCCTACT 629
Qy 336 CACTGAGGCGACCTCAAGTTTGAATCTCGATCGAAGACGGCCCGGACAGAGCAAGATAAC 395
Db 628 TACTGAATCTGATTTTGAATTCGATCTCGATAGAGGAGGCGCTGTCAGAGCAGGATCAC 569

396	Qy	CTCCACAGGAAGGACATGATACATGCAGAAATCTCTGCAGGACACATGAGGGGTGAC	455
439	Qy	CTCCACAGGAAGGACATGATACATGCAGAAATCTCTGCAGGACACATGAGGGGTGAC	498
568	Db	AATCTCTAGAAAGAGACATGATCTTGGCGGATATCGCTGAGATCTCTGAAGGAATGAC	509
456	Qy	TACTGGAAGCCCAATCTTTGTTATTTGTCTCCCAACACAGATCAAAATAGGCAGTGCATCACCG	515
508	Db	GACAGGAACACCTATCCATGTGTTGTACCAAAACACAGATCAGAGAGGACTTCGTTTACAG	449
516	Qy	TGAAATAGCCATGTGTACCGACCTTCTCATGCGACGCAACTTATGACTTCAAGTACGG	575
448	Db	TGAAATGTCCGGTTGCCATATAGACCCATCCGATCTGTATGCAACTTATGACATGAAGTATGC	389
576	Qy	TGTTTAGAGCTGTACAGGGAGGTGGGAGGTCTCTCGGCGAGAAAAACCGTTCGAAAGCGGTGGC	635
388	Db	TGTCAGATCAGTGCAGGGTGGAGGAAGATCTTCAGCTAGAGAGACCATTGGAAGAGTTGC	329
636	Qy	TGCAGGGGGCTCCCGAGAAATTTCTTAAGTCAAAATGTGAGATCTTTGTCGTT	695
328	Db	TCCTGAGCTTTGGGCGAGAGAATTTTGAAGCAATTTGCAGGACCTGAGAATCTTCGCCTA	269
696	Qy	TGTTTCCAAAGTGCATCAGGTTGTGCTCCCGAGAGAGCGCGTTTCGATTATCGGGTCTGTAAC	755
268	Db	TGTCTCGAAGTTCACCAATGTGTACTTTCAGAGAGAATTTGGTAGACAGAGAATTTTACAC	209
756	Qy	TTTGTGAACAGATAGAGAGCAACATCGTTTAGATCTCTGATCCAGAGTACGACAGAGAAT	815
208	Db	TCGG--ACAGATAGAAAAATAAATTTGTTCAGAAGCCCTAAATCCGAAATAATCCGAAAAGA	151
816	Qy	GATAGACGCAATCGACAGAGTACGAGTTTCGAGGGGATTCGGTCGGTGGAGTGCATCATG	875
150	Db	GATTACTGCGAATGAATCCGTCAGAGAGCATGAAGAAATCTTTGTGTGGGTTTTGACCTG	91
876	Qy	CGTCCGTAGAAA	887
90	Db	CATTGTCCGAAA	79

```

RESULT 6
US-09-610-040-7
; Sequence 7, Application US/09610040
; Patent No. 645217
; GENERAL INFORMATION:
; APPLICANT: Boyes, Douglas
; APPLICANT: Davis, Keith
; APPLICANT: Woessner, Jeffrey
; APPLICANT: Grolach, Jorn
; APPLICANT: Hamilton, Carol
; APPLICANT: Hoffman, Neil
; APPLICANT: Kloti, Andreas
; APPLICANT: Zayed, Adel
; APPLICANT: Alicenzi, Robert
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHAS
; TITLE OF INVENTION: CHORISMATE MUTASE EXPRESSION OR ACTIVITY IN PLANTS
; FILE REFERENCE: 9128.14
; CURRENT APPLICATION NUMBER: US/09/610,040
; CURRENT FILING DATE: 2000-07-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: "n" indicated any nucleotide (A, C, T, or G).
; NAME/KEY: misc_feature
; LOCATION: (20)..(21)
; OTHER INFORMATION: "n" indicated any nucleotide (A, C, T, or G).
; NAME/KEY: misc_feature
; LOCATION: (29)..(29)
; OTHER INFORMATION: "n" indicated any nucleotide (A, C, T, or G).
; NAME/KEY: misc_feature

```

```
; TITLE OF INVENTION: EXPRESSION OR ACTIVITY IN PLANTS
; FILE REFERENCE: 2035DIV1
; CURRENT APPLICATION NUMBER: US/10/267,763
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 09/610,040
; PRIOR FILING DATE: 2000-07-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Arabidopsis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1053)
; OTHER INFORMATION: n = A,T,C or G
US-10-267-763-1

Query Match      16.9%; Score 274; DB 4; Length 1053;
Best Local Similarity 64.0%; Pred. No. 6.1e-78;
Matches 430; Conservative 0; Mismatches 240; Indels 2; Gaps 1;

Qy 216 CCTAGAGTGAGGCGATCTGGAACACGTTTGGAACTACTTTGAGGTTCAGGTTCGAACCTATGG 275
Db 748 CTTCCAGATACAGCTACTGGAAGTTTCATATGGGACTCATTTTCGAGTTTCAACTTTGG 689
Qy 276 TGAATCTCATGGGGTGTGGTGTGGTATCATAGTGTGTTCACCTAGATTCCTACT 335
Db 688 AGAATCATCGAGGAGGAGTGGTGTGATCATTTGATGTTGTCTCTCGTATTCCTACT 629
Qy 336 CACTGAGGCGAGCTTACAGTGTGAAGTTCGATCGATGAGAGCGGCCGACAGAGCAAGATAAC 395
Db 628 TACTGATCTGATTTGCAATTCGATCTCGATAGAGGAGGCTGTGTAGAGCAGATCAC 569
Qy 396 CTCCACAAGGAGGAGCTGATACATGCAAAATTCCTGTCAGGACACATCAATAGGAGTATCAC 455
Db 568 AACTCCTAGAAAAGAGAGCTGATCTTCCGGATATCGTCTGGAGTCTCTGAAGGAATGAC 509
Qy 456 TACTGGAAGCCAAATCTGTTATTTGTCGCAACACACAGATCAATAGGAGTATCACCG 515
Db 508 GACAGGAACACCTATCCATGTTGTTGTTACCAACACACAGATCAGAGAGGACTTGATACAG 449
Qy 516 TGAATAGCCAAATGTATCCGACCTTCTCATGACAGACGCACTTATGACTTCAAGTACGG 575
Db 448 TGAATGTGCGTGTGCTATAGACCATCGATGCTGTGCACTTATGACATGAGTATGG 389
Qy 576 TGTAGAGCTGTACAGGAGGTGGAGGTCTCTCGGCGAAGAAACCCGTTGGAAGGGTGGC 635
Db 388 TGTGAGATCAGTGCAGGGTGGAGGAAGATCTTCAGCTAGAGAGACCAATTGGAAGAGTTGC 329
Qy 636 TCCAGGGGCCCTCCCAAGAAATTTTAAGCTCAAAATGTGGATTAGAGATCTTGTGCTT 695
Db 328 TCTTGGAGCTTTGGGCAAGAGAAATTTTGAAGCAATTTTCAGGAACCTGGAATCTTGGCCTA 269
Qy 696 TGTTTCCAAAGTCATCAGTGTGCTCCAGAGACGCGTGTGATTTATGCTGTGTAAC 755
Db 268 TGTCTCGCAAGTTCACCAATGTGTACTTCCAGAGAAATTTGTAGACAGAAATTTACAC 209
Qy 756 TTTGAAACAGATPAGAGACCAATCGTTAGATGTCTCTGATCCAGAGTACGAGAGAGAT 815
Db 208 TCCG--ACAGATAGAAATAACATTTGTGTCAGAGCCCTAAATCCGAAATAATCCGAAAGA 151
Qy 816 GATAGACCAATCGACAGATGACGAGTTCGAGGGGATTCGGTCTGGTGTGATGATCAGATG 875
Db 150 GATAACTCGGAATGAATCCGTCAGAGCATAGGAAATCTTGTGTTGGTGGGTTTGGACCTG 91
Qy 876 CGTCTGCTAGAA 887
Db 90 CATTGTCGAAA 79

; Sequence 7, Application US/10267763
; Patent No. 6800459
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: Boyes, Douglas
; APPLICANT: Davis, Keith
; APPLICANT: Hamilton, Carol
; APPLICANT: Hoffman, Neil
; APPLICANT: Kloti, Andreas
; APPLICANT: Woessner, Jeffrey
; APPLICANT: Zayed, Adel
; APPLICANT: Ascenzi, Robert
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: MODULATION OF CHORISMATE SYNTHASE AND CHORISMATE MUTASE
; TITLE OF INVENTION: EXPRESSION OR ACTIVITY IN PLANTS
; FILE REFERENCE: 2035DIV1
; CURRENT APPLICATION NUMBER: US/10/267,763
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 09/610,040
; PRIOR FILING DATE: 2000-07-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1053)
; OTHER INFORMATION: n = A,T,C or G
US-10-267-763-7

Query Match      16.9%; Score 274; DB 4; Length 1053;
Best Local Similarity 64.0%; Pred. No. 6.1e-78;
Matches 430; Conservative 0; Mismatches 240; Indels 2; Gaps 1;

Qy 216 CCTAGAGTGAGGCGATCTGGAACACGTTTGGAACTACTTTGAGGTTCAGGTTCGAACCTATGG 275
Db 306 CTTCCAGATACAGCTACTGGAAGTTTCATATGGGACTCATTTTCGAGTTTCAACTTTGG 365
Qy 276 TGAATCTCATGGGGTGTGGTGTGGTATCATAGTGTGTTCACCTAGATTCCTACT 335
Db 366 AGAATCATCGAGGAGGAGTGGTGTGATCATTTGATGTTGTCTCTCGTATTCCTACT 425
Qy 336 CACTGAGGCGAGCTTACAGTGTGAAGTTCGATCGATGAGAGCGGCCGACAGAGCAAGATAAC 395
Db 426 TACTGAACTCTGATTTGCAATTCGATCTCGATAGAGGAGGCTGTGTGAGAGGAGTAC 485
Qy 396 CTCCACAAGGAAGAGAGCTGATACATGCAAAATTTCTGTGAGGACACATGAAAGGGGTGAC 455
Db 486 AACTCCTAGAAAAGAGAGCTGATCTTGGCGGATATCGTCTGGAGTCTCTGAAGGAATGAC 545
Qy 456 TACTGGAACCCAAATCTTGTATTTGTTCCAAACACAGATCAAAATAGGCGAGTATCACCG 515
Db 546 GACAGGAACACCTATCCATGTTGTTGTTACCAACACACAGATCAGAGAGGACTTGTATTACAG 605
Qy 516 TGAATAGCCAAATGTGTACCGACCTTCTCATGACAGACGCACTTATGACTTCAAGTACGG 575
Db 606 TGAATCTCGGTTTGGCTTATAGACCATCGCATGTGTGTCGCACTTATGACATGAAGTATGG 665
Qy 576 TGTTAGAGCTGTACAGGAGGTGGAGGTCTCTCGGCGAAGAAACCCGTTGGAAGGGTGGC 635
Db 666 TGTGAGATCAGTGCAGGGTGGAGGAAGATCTTCAGCTAGAGAGACCATTTGGAAGTGTG 725
Qy 636 TGCAGGGGCCCTCCCAAGAAATTTTAAAGTCAAAATGTGGATTAGAGATCTTGTGCTT 695
Db 726 TCCTGGAGCTTTGGGCAAGAGAAATTTTGAAGCAATTTTCAGGAACCTGAGAATCTTGCCTA 785
Qy 696 TGTTCCAAAGTCATCAGGTGTGCTCCAGAGAGACGCGTGTGATTTATGCTGTGTAAC 755
Db 786 TGTCTCGCAAGTTCACCAATGTGTACTTCCAGAGAAATTTGTAGACACGAGAAATTTACAC 845
Qy 756 TTTGAAACAGATAGAGAGCAACATCGTTAGATGTCTGATCTCAGAGTACGAGAGAGAT 815
```

846 TCCG--ACAGATAGAAAATAACATTGTGAGAAGCCCTAAATCCGAATAATCCGAAAAAGA 903
Db . : | ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
816 GATAGACGCAATCCGACAGACTACGAGTTCGAGGGCGAATCCGTCGGTGAGTGATCACATG 875
QY
904 GATAACTCGGAATGAATCCGTCAAGACGATTAAGGAATCTTTGTTGGTGGTTTTGCACCTG 963
Db
876 CGTCGCTAGAAA 887
QY
964 CATTGTCGGA 975
Db

RESIN, T 9

```

US-09-248-796A-4099
; Sequence 4099, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 4099
; LENGTH: 1080
; TYPE: DNA
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (36)
; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unknown
; US-09-248-796A-4099

```

Query Match	12.5%;	Score 203.8;	DB 4;	Length 1080;
Best Local Similarity	52.1%;	Pred. No. 4.1e-55;		
Matches 541;	Conservative	0;	Mismatches 473;	
			Indels 24;	Gaps 3;

238	GGTGTGTTATCAGTGGTTGTCACCTAGAAATTCCTACTCAGGAGGACCTCAAGTT	357
7	GCCTTTCTTTCCCTGGTGGCAGAGTCGCGATGTCATTAACTGAAAGCGGATATTCACCA	66
358	GAATCGATCGAAGACGCGCCGGACAGAGCAGAGTAACCTCCACAAGGAAGAGACTGAT	417
67	CAATTGACTCGTATACGTCAGGTCAAAGTAATATTCTACCCGAGAGAGATGMAAAGAT	126
418	ACATGCAAAATTCGTTCAGGGACACATGAAGGGGTGACTACTGGAAACGCCAAATCTTTGTT	477
127	TTATGTTGAAATACAAAGTGGAATCGAGAAATGGATTAACTTTGGGTTCACTATTGGAATG	186
478	ATTGTGCCAARACACAGATCAAAATAGGACGTGATCACCGTGAANATGCCAATGTGTACCGA	537
187	ATTGTGAAATAAAGATCATCGTCTGGTGATTATAGTGAGACCGGATTTGTATCCAAAG	246
538	CGTTCTCATGTCAGACGCAACTTATGACTTCAAGTACCGTGTTAGAGCTGTACAGGGAGGT	597
247	CCAACTCATGCGACTGGACCTATATACAGAAATACGGAACCAAACTAGTTCTTGGTGGT	306
598	GGAGGTCTCTCGGGCAGMAAAACCGTTTGGAAAGGGTGGCTGCGAGGGGCCCTCCCAAGAAA	657
307	GGTAGATCATCTGCTAGAGAAACTATTTGGTAGAGTTTCCGCGAGGAGCCATTGCGAAGAA	366
658	ATTCTTAAGCTCAAAATGTGGATTAGAGATCTTGTGCTTTGTTTTCCTCAAAAGTGATCAGGT	716
367	ATTTTGGCCAAGTGAATATGTTGAAATTTGTTGCCCTTTGTAGTGTCTATTGGTGAATA	426
717	-----TGTGCTCCCAAGAGACCGCGTTGATTATCGGGTCTGTAACTTTTG	759
427	TCTATGAATAAATCTCCTCAAGATGCAAAATTCCAAGAACTTTTAAACAACATCACTAGA	486

Qy	760	GAACAGATGAGAGCAACATCG---	TTAGATGCTCGATCCAGAGTACGCAGAGAGATG	816
Db	487	GAACAGCTCGATGGTGTAGGTCCAA	TAAAGATGTCAGATCGAAATGTC	546
Qy	817	ATAGAGCCAATGACACAGAGTACGAG	TTCCGAGGGGATTCGGTCCGTGCGAGTGATC	876
Db	547	GTTTAAAGTGATGAAATAATCGTGAT	GCTAAAGATTCCAATGGTGGTGTGTCAC	606
Qy	877	GTCGCTAGAAACGTTCCCTCGCGGCT	CGGTTCTCTGTGTTCCAGCAAGCTCGAAT	936
Db	607	GTTTATCAGAACTGTCCTCAATTCGAT	TAGGTGAACCATGTTTTCGATAAATGGA	666
Qy	937	CTCGCAAAAGCTATGCTTTCTTCTAT	TCTCGCAGCAACCGGTTTCAGATTCGC	996
Db	667	TTGGCTCATGCCATGTTATCATTTAC	CCCGCTACCAAGGGGTTGAAATTTGGT	726
Qy	997	GCGCGGACCGACTTGACAGGAAGTG	AGCAATAATGATGAGTT---TTATATGG	1053
Db	727	GAAGGTATCAAGATCCGTGGTTCA	AAACATACGATGCATTTATTTATGAT	786
Qy	1054	GGAAGTGTCCAGGACCGGACTAATCG	CTCGGTGGTGTGCAGGAGGGAATATCG	1113
Db	787	GGAAGATTAAAGAACTGAAACAAAC	ATAGTGGTGGTATCCAAAGGAGAA	846
Qy	1114	GAGATTGTGCATTTCAAAGTTGCTTT	TAAAGCCGACACCATCTATCGGGGTGA	1173
Db	847	GAAAAACATTTATTTCTCAGTTGCT	TTCAAATCGGCTGCTACTATCAGTCA	906
Qy	1174	ACCGTGTCAAGGGAGCGGTACGACGT	TGAGCTTCTAGCAAGGGCGCCATGAC	1233
Db	907	ACTGCTACTTACATGGTAAAGTG	TGGTGTGCTAGAGGTAGACACGAT	966
Qy	1234	GTCGCCCTTCGAGCTGTTCTCTGTG	TGGGAATCCATGGCCCGTTGGTCTC	1293
Db	967	GTCACCCCAAGACAGTTCCAA	TGTTTGAAGCCATGACTGCATTAG	1026
Qy	1294	CTGATGGCGCACGTGGCT	1311	
Db	1027	TATATGATTTCAACAAGCT	1044	

RESIST 10

```

US-09-313-294A-4734
; Sequence 4734, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 4734
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700348949H1
; NAME/KEY: unsure
; LOCATION: 44, 54
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-4734

```

Query Match 11.4%; Score 185; DB 4; Length 300;
Best Local Similarity 84.3%; Pred. No. 2.1e-49;
Matches 231; Conservative 0; Mismatches 41; Indels 2; Gaps 2;
QV 546 TGCAGACGCAACTTATGACTTCAGTACGGTGTAGAGCTGTGTA-CAGCGAGGTTGGAGGT 604

Db 1 TGCAGATGCAACTATGACTTCAAGTATGAGTGTAGAGCTGTANACGGGACACNGAAGT 60
Qy 605 CCTCGGCGAGAAAAACCGTTGAGAGGGTGGCTGCGAGGGGCCCTCCCAAGAAAAATCTTAA 664
Db 61 CATCAGCGAGAGAAACCAATTTGGCAGGGTGGCTGCGAGGAGCTCTTGCAGAAAGAAAAATCTAA 120
Qy 665 AGCTCAAAATGTGGATTAGAGATCTTGTCTGTTTGTTCCTCAAAAGTGCATCAGGTGTGCTCC 724
Db 121 AGCTCAAAATCAGAGTGGAGATCTTGGCATTTTGTCTTAAAGTGCAACAGTGTGACTTC 180
Qy 725 CAGAGACCGCGTTGATTATGGGCTGTAACTTTGGAAACAGATAGAGCAACATCGTTA 784
Db 181 CAGAGATGCGAGTTGATTATGAGACTGTAACTTTGGAAACATGAGAGCAACATCGTTA 239
Qy 785 GATGCTCTGTATCCAGAGTACCGAGAGAGATGAT 818
Db 240 GATGCTCTGTATCCAGAGTATGAGAGAGATGTT 273

RESULT 11

US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENORU, SHUJI
; APPLICANT: KATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 11.1%; Score 180.2; DB 4; Length 640681;
Best Local Similarity 50.1%; Pred. No. 2.2e-45;
Matches 532; Conservative 0; Mismatches 493; Indels 36; Gaps 2;
Qy 232 TCTGAAACACGTTTGGAACTACTTTCCAGGTTGCAACTATGGTGAATCTCATGGGGT 291
Db 102985 TCTGAAATACAAATTTGGGAAAAATATCTGTGTAACCACTTTTGGCGAGTCACCGAGAA 102926
Qy 292 GGTGTGTTGTTATCAGTGGTTGTCCACTAGAAATCCACTCACTGAGGCGAGACCTA 351
Db 102925 GCATAGGGTGCATAATTTGACGGAACACCTCTCGTCTTGAATTTATCTTGTAAAGATTG 102866
Qy 352 CAAGTTGACCTCGAGAGCGGCCGAGCAGACAGCAATAAACCCTCACAAGGAGAG 411
Db 102865 CAATAGATTGAAATCGTAGAAGACAGACTCTCCCGTTACAGACTTTGCGCGTGAA 102806
Qy 412 ACTGATACATGCAAAATTTCTGTGAGGAGACATGAAAGGGGTGACTACTTGGAAACGCGCAAT 471
Db 102805 CCTGATGAAGTTAATATATCTTCCGGTATATTTAAGCGGGTCACAACTGGTACTAGTATT 102746
Qy 472 CTGTGTTATTTGCCAAACACAGATCAATAGGCACTGATCCCGTGAATAATAGCCATGTG 531
Db 102745 GGTTTAAATTAATTAATCATGATCATAGGCTCAAGACTATAGCGATATAAAGAAATTTA 102686
Qy 532 TACCGACCTTCTCATGCGAGCGCAACTTATGACTTCAAGTAGCGGTGTAGAGCTGTACAG 591
Db 102685 TTTCACCGGNCATGCCGATTATCTTATGAAAAAATATGGAATTTAGAGATTATCGC 102626
Qy 592 GGAGGTGGAGGTCTCTCGGCGAGAAAAACCGTTGGAAGGGTGGCTGCGAGGGGCCCTCCCC 651

Db 102625 GGGGAGGTAGATCTTCTGCTCGCAAACTGCTATGAGAGTTGCGAGCAGGCTATAGCA 102566
Qy 652 AAGAAATTTCTAAGCTCAAAATGTGGATTAGAGATCTTGTGCGTTGTTTCCAAAGTGCAT 711
Db 102565 AAAAAATACCTTAATGAGAAATATGGGATAACTATTTCGACATATTTGTGAGCAATGGC 102506
Qy 712 CAGGTGTGCTCCAGAGACGCGGTTGATTATATGGGTCTCTAACTTTGGAAACAGATAGAG 771
Db 102505 AATATCAATGCCCT-----TTTAAATCTTGGCAAGAGTCGAA 102467
Qy 772 AGCAACATCTGTAGATGCTCTGATCCAGATAGCGAGAGAGATGATAGACGCAATCGAC 831
Db 102466 AATAATCCTTTTCTGCTGATCCTGAAAAAATTTAGCACTAGAAAAATTTAAATTTAAA 102407
Qy 832 AGAGTACAGTTCGAGGGGATTCGGTCGGTGGATGATCATCGCTCGCTAGAAACGTT 891
Db 102406 TATCTCAAAAAAATAGGTGATTCGATTGGTGGCTGGAATAACAATTTTGTGAAAAATATC 102347
Qy 892 CCTCGCGGGCTCGGTTCTCTGTTTCGACAAGCTCGAAATCCGAACTCGGCAAAAGCTATG 951
Db 102346 CTTGTAGACTTGGAGAACCACTTTTGTGATCGTCTTGTGATTTATCACGCGCTTG 102287
Qy 952 CTTTCTATTCTCGAGCAACGGGTTTCGAGATTGGCAGCGGATTCGCGGAGCCGACTTG 1011
Db 102286 ATGAGTATTAAATGCGCAAAAGGAGTAGAAATTTGGAGACGGTTTTCAGTAATAAATCAA 102227
Qy 1012 ACAGGAATGAGCANTATGATGATGTTTTATATGATAAGGCTGGAAGTGTGAGGACACGG 1071
Db 102226 CGAGGAAGTGAACATCGTGATGAAATTTACGCCAAGGATTTTAA----- 102182
Qy 1072 ACTAATCGCTCGGTTGGTGTGCGAGGAGGATATCGAATGTGTGAGATTGTGCACCTTCAAA 1131
Db 102181 ACTAATCATCTCGTGGTGTATCTAGTGGTGTATAGTAAATGGTTCGAGAAATTTGACTAAA 102122
Qy 1132 GTTGCTTTTAAAGCGACACCATCTATCGGGGTGAAACAGAACACCGTGTCAAGGGAGCGT 1191
Db 102121 GTAGCATTTAAACCTACATCAAGTATTTCGAAAAAGCAGGCAATACATAAATAAATAAAC 102062
Qy 1192 CAGAACGTTGAGCTTCTAGCAAGAGGGCGCCATGACCCATCGCTCGCCCTCGAGCTGTT 1251
Db 102061 GAAAAAGTTCAAAATAGTTTACTAAAGGACGACATGATCCATGTGTAGTGTTCACGCGCTGT 102002
Qy 1252 CCTGTGTGGAATCCATGCGCGGCTTGGTCTCTCATGACCA 1292
Db 102001 CCAATCACTGAGCAATGTAGCAATTTGATTAATGATCA 101961

RESULT 12

US-09-543-681A-120
; Sequence 120, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 120
; LENGTH: 1089
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-120

Query Match 11.0%; Score 179.6; DB 4; Length 1089;
Best Local Similarity 49.8%; Pred. No. 3.1e-47;
Matches 541; Conservative 0; Mismatches 509; Indels 36; Gaps 2;
Qy 233 CTGGAACACGTTTGGAACTACTTTCCAGGTTCGAACTATGGTGAATCTCATCGGGGTG 292

Db 8 CGGAAACAGTATCGGCAATTTATTAGAGTAACCACTTTTGGTGAGTCTCATGCGACAG 67
QY 293 GTGTTGGTGTGTTATCAGTGGTGTCTCACTAGAAATCCACTCAGTGGACGACCTAC 352
Db 68 CGTTAGGTTGCAATTTGGTGTCTCTCCGGAAGTACCTTTAAACGCAAGCGATCTAC 127
QY 353 AAGTTGAACCTCGATCGAAGACCGCCCGGACAGACAGCAATAAATCCCTCACAGGAAGGAGA 412
Db 128 AAGTTGATTTAGATAGACGTAACCGGGAATCTTACGTTATACACAAACGTTAGAGGC 187
QY 413 CTGATACATGCAAAATTTCTGTGAGGACACATGAAGGGGTGACTTCTGGAACGCCAATTC 472
Db 188 CTGATCAAGTGCATTTTATTCGGGTGTGTTTAAATGCTGTAAACAACGGGAACCAATTTG 247
QY 473 TTGTTATTGTCCTCCAAACACAGATCAATAGCGAGTATCACTGATGATGAAATAGCAATGTGT 532
Db 248 GATTTATTAAAGAAATACCGATCAGCGCTCTCAAGATTTATGCGAAATTAAGATGTAT 307
QY 533 ACCGACCTTCTATGACAGCGCAACTTATGACTTTCAAGTACGGTGTAGAGCTGTACAGG 592
Db 308 TCCGCCCCAGGCGATCGACATACACCTATGAACAGAAATATGTTTACGTGATATCGTG 367
QY 593 GAGTGGAGGCTCTCGGCGAGAAACCGTTGGAAGGGTGTGACAGGGGCCCTCCCA 652
Db 368 GCGGGGACCGCTCTTCTGTCGAGAAACCGCCATGCGTGTGCGAGCGGTGCTATTGCCA 427
QY 653 AGAAATTTCTTAAGCTCAAAATGTGGATTAGAGATCTTGTGTTTGTTCAAAGTGCATC 712
Db 428 AAAATACCTTAACAAAAATTTGGTATTGAGTAAAGGTTACTTATCTCAATT----- 482
QY 713 AGTTGTGCTCCAGAAAGCGGGTTGATTATGGGTCTGTAACTTTTGGACAGATAGAGA 772
Db 483 -----AGGCGCGATTAGTTGTGAGTTAGTTGTTGTTGTTTCAATT----- 482
QY 773 GCACATCGTTAGATGCTCTGATCCAGATAGACGAGAGAGATGATAGCGCAATCGACA 832
Db 527 CCACCCATCTTTTCCCGAGATCTCTCTCGCTAGATGCCCTTGATGAATATATGCGAG 586
QY 833 GAGTACGAGTTCCGAGGGGATTCGTCGGTGGAGTGATCATCGCTCGCTAGAAAGTTC 892
Db 587 CGCTAAAAAAGAGGTAAATCTATAGTGGCAAGATTTACCGTGGTTGCAAGAGGTGAC 646
QY 893 CTGCGGGTCTGGTCTCTGTTGTCGCAAGCTCGAATCCGAATCGCAAAAGCTATGC 952
Db 647 CTGCAAGATTAGTGAACCCGTTTGTGATAGACTTGATCGCGATTTGGCTCATGCGTTAA 706
QY 953 TTCTATTCTCCGAGCAGCGGTTGCGAATTCGAGATTGCGAGCGGATTCGCGGGACCGACTTGA 1012
Db 707 TGAGCATCAACGAGTAAAGCGGTGAAATCGAGATGGTTTGTGATGTTGACCTTAA 766
QY 1013 CAGGAAGTGAGCAATATGATGAGTTTATATGATAAGCTGGAAGTGTGAGGACACGGA 1072
Db 767 AAGGACTGAACCCCGATGAATTTACCAAGAGGATTAGCAG----- 812
QY 1073 CTAATCGCTCGGTTGTCGAGGAGGATATCGAATGTTGAGATTTGCACTTCAAAG 1132
Db 813 -TAATCAGCGGGGGCGGTATTAGCGGCATTTAGTGGTCAACCTATTATTGCACACA 871
QY 1133 TTGCTTTTAAAGCCGACACCATCTATCGGGGTGAACAGACACCGTGTCAAGGGGCGTC 1192
Db 872 TTGCTTTTAAAGCCACTTCAAGTATTACTATGCTGGTGAACACTTAATCGAAGGTG 931
QY 1193 AGAAGTTGAGTCTTACAGAGAGCGGCCATGACCCATGCGTCCGCCCTCGAGCTGTC 1252
Db 932 AAGAGTGCATATGATCAACCAAGGGCGACATGATCCGTTGTCGGTATCCGACTGTC 991
QY 1253 CTGTTGTTGGAATCCATCGCGGTTGTCCTCATGACCGACCTGATGCGGCACTGTC 1312
Db 992 CTATTGCAAGAGCTATGATGCTATTGTTACTTGTATGATCATGCTGCTCAACCGGGTC 1051
QY 1313 AGTGGC 1318
Db 1052 AATGTG 1057

RESULT 13

US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6503729
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)

Db 357920 AGATTGGAAGATTGAAGGAGATTTTCAGCTACTATAAAAATCCAGAGGTTTTTGAATAATG 357979
QY 752 TAACTTTGGAACAGATAGAGCAACATCGTTAGATGTC-----TGATCCAG 799
Db 357980 AAAAATCCTTAGAGAGATTATAGAGATTATTGAAAGATTATCCATTGAGATGTCATCAA 358039
QY 800 AGTACACAGAGAAGATGATAGACGCAATACGACAGATACGAGTTGAGGGGATTCGGTCG 859
Db 358040 TGAATGAGAAGAGATGGAGGAGTATGTTTAAAGCAATGGAATAAAGATAGTGGTTG 358099
QY 860 GTGAGTATCATACGTCGCTAGAAAAGTTCTCGCGGCTCGGTTCTCTGTTGTTCCG 919
Db 358100 GAGGAGTTTGAATTTGTTGCAATTAATTTCTCTGTTGGAGTTGGAATCCAATATTTCA 358159
QY 920 ACAAGCTCGAATCCGAATCGCAAAAGCTATGCTTTTCTATTCTCGCGACCAAGGTTCCG 979
Db 358160 ATAGTTTAAATGAGAAATTTGGCAAGAGCTTTTAATGAGTATAAATGCTGTTAAAGGAGTTG 358219
QY 980 AGATTGGCAGCGGATTCGCGGACCGACTTCACAGGAAGTGAGCATAATGATGATTTT 1039
Db 358220 AGATAGGCGCTGGTTTAAAGCGCTGAGATGTATGGAAGTGAGATGAACGATGATGT 358279
QY 1040 ATATGATGAAGCTGGAAGTGTGAGACAGGACTTAATCGCTCGGTTGTTGCGAGGGAG 1099
Db 358280 ATTTTATGACGACAAAAATATAAGATTCAAAACAACTCGCTGGCATATTGGGAG 358339
QY 1100 GGATATCGAATGTGAGATTGTGCACTTCAAAAGTTGCTTTTAAGCGCACACCATCTATCG 1159
Db 358340 GAATTAGCTGTGAACTCAATAGTTTAAAGAAATGCAAGCAACACCTTCAATAG 358399
QY 1160 GGGTGAACAGAACCCGT---GTCAAGGAGCGTCAGAACGTTGAGCTTTAGCAAGAG 1216
Db 358400 GTAAAAAGCAAAACCAATAAATTTAAACCTTAGAAATGTTGAAATGAAATGAAG 358459
QY 1217 GGGCGCATGACCCATCGCTGCGCCCTCGAGCTGTTCTGTGTGGAATCCATGCGCGCT 1276
Db 358460 GAAGACACGACCGAGTTATAGTTTCAAGGATGTTCCAGTGCTGAAGCAATGTTGCTA 358519
QY 1277 TGGTCTCATGACCAAGCTGAT 1298
Db 358520 TAACTTAGCTGATTTGATGAT 358541

RESULT 14
US-09-692-570-1
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Buit et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus jannaschii
; Patent No. 6797466
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:

NAME/KEY: misc_feature	
LOCATION: (309418)..(309418)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (312837)..(312837)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (312993)..(312993)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (319226)..(319226)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (559167)..(559167)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (559241)..(559241)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (600992)..(600992)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (622708)..(622708)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (657081)..(657081)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (657203)..(657203)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (674435)..(674435)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (682142)..(682142)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (713652)..(713652)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (741684)..(741684)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (779155)..(779155)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (779676)..(779676)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (855539)..(855539)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (871619)..(871619)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	

:	LOCATION:	(1084830)..(1084830)
:	OTHER INFORMATION:	n equals a, t, c, or g
:	FEATURE:	
:	NAME/KEY:	misc feature
:	LOCATION:	(1096846)..(1096846)
:	OTHER INFORMATION:	n equals a, t, c, or g
:	FEATURE:	
:	NAME/KEY:	misc feature
:	LOCATION:	(1119881)..(1119881)
:	OTHER INFORMATION:	n equals a, t, c, or g
:	FEATURE:	
:	NAME/KEY:	misc feature
:	LOCATION:	(1130881)..(1130881)
:	OTHER INFORMATION:	n equals a, t, c, or g
:	FEATURE:	
:	NAME/KEY:	misc feature
:	LOCATION:	(1310988)..(1310988)
:	OTHER INFORMATION:	n equals a, t, c, or g
:	FEATURE:	
:	NAME/KEY:	misc feature
:	LOCATION:	(1313224)..(1313224)
:	OTHER INFORMATION:	n equals a, t, c, or g
:	FEATURE:	
:	NAME/KEY:	misc feature
:	LOCATION:	(1349473)..(1349473)
:	OTHER INFORMATION:	n equals a, t, c, or g
:	Query Match	10.9%; Score 178; DB 4; Length 1664976;
:	Best Local Similarity	49.9%; Pred. No. 2.3e-44;
:	Matches	550; Conservative 0; Mismatches 525; Indels 27; Gaps 3;
Qy	224	TGAAGGCATCTGAAAACACCGTTTGGAACACTCTTTCAGGTGGCAACCATTATGGTGAATC TC 283 357440 TGGTGACATTAATGAACACCTATCGGGATATGTTTTAGAGTTTACAGTTTTTGGAGAAAGTC 357499
Dd		
Qy	284	ATGGGGGTGGTGGTTGTGTTATCAAGTGTGTGCACCTAGAAATCCCACTCACTGAGG 343 357500 ATGGAAAGGCTGTTGGAGCAGTTGTTGATGGATGCCAGCTAATCTGCTTTTATCTGAAG 357559
Dd		
Qy	344	CAGACCTACAAGTTGAATCGATCGAAGACGGCCGGACAGAGAATAAACCTCCACA 403 357560 AGGNATFCCAANAAGAGCTTGACAGGAGAGACCGAGGAGAGCATTTCTCAAACACCA 357619
Dd		
Qy	404	GGAAGGAGACTGATACATGCAAAATCTGTCAAGGACACATGAAGGGTGACTACTGGAA 463 357620 GAAGAAGAGGATAAAGTTGAATCTTATCAGGAATTTTTGAGGGAAAACTACTGGAG 357679
Dd		
Qy	464	CGCCAAATCTTTGTTATTTGTCCEAAACACAGATCAAAATAGGCAGTGATCACCGTGAATAG 523 357680 CTCCTATTTGCTCAATAGTCTTATACCAAAAACATGAGACCTTAAAGATTACTCAAAAATTA 357739
Dd		
Qy	524	CCATGTGTACCGACCTTCTCATGCAGACGCAACTTATGACTTCAAGTACGGTGTAGAG 583 357740 AAGNATACCAAGACCTGGACATGCAGATTTAACTTAGATTTGAAGTATAAAACACTATG 357799
Dd		
Qy	584	CTGTACAGGAGGTGGCAGGTCTTCGGGCAGAAAAACCGTTTGAAGGGTGGCTGCAGGGG 643 357800 ATTATAGGGNGAGGAAGGCAAGTGGTAGAGTTAACGATAGGCATGTTATTGGAGGAG 357859
Dd		
Qy	644	CCCTCCCAGAAAAATTTCTTAAGCTCAAAATGTGGATTAGAGATCTTGTCTGTTGTTTCCA 703 357860 CTATTGCTCAAAAAGCTTCTATCTTACACATACAAAAAATTTATTGTTTATACCATAA 357919
Dd		
Qy	704	AAGTGCATCAGTTGTG-----CTCCCCAGAGACCGGTTGATTATGGGTCTG 751 357920 AGATTGGNAAGATTGAAGGAGATTTCAGCTACTATAAAATCCAGAGGTTTTTGGAAATG 357979
Dd		
Qy	752	TAACTTTGGAAACAGATAGAGCAACATCGTTTAGATGTCC-----TGATCCAG 799 357980 AAAAATCCTTAGAGAGATTATAGAGATTATTGAAGATTATTGAAGTAAATCCATTGAGATGTCATCAA 358039
Dd		
Qy	800	AGTACGAGAGAAGATGATAGACGAATCGACAGAGTACGAGTTTCGAGGGGATTCGGTCG 859

Db 358040 TGAATGAGAAAGAGATGAGGAGTATGTTTTAAAGGCAATGGAATAAAGATAGTTGT 358099
Qy 860 GTGAGATGATCACATGCTGCTAGAAACGTTCTTCGGGCTCGGTTCCTCTGTGTTCG 919
Db 358100 GAGAGTTGTTGAATTTGTCATTAATGTTCTGTTGGAGTTTGGAAATCCAATATCA 358159
Qy 920 ACAAGCTCGAATCCGAATCGCAAAAGCTATGCTTTCTATTCCTCGGAGCAACGGGTTG 979
Db 358160 ATAAGTTAAATGGAGAAATGGCAGAGCTTTAATGATATAAATGCTGTTAAAGGAGTTG 358219
Qy 980 AGATTGGCAGCGGATTCGCGGGACCGACTTGACAGAACTGAGCAATAATGATGAGTTT 1039
Db 358220 AGATAGGGCTGTTGTTTTAAAGCGCTGAGATGATGAAAGTGAATGAGATGAACGATGAGATG 358279
Qy 1040 ATATGGATAAGGCTGGAAGTGCAGGACCGGACTATCCCTCGGTGGTGTGCAGGAG 1099
Db 358280 ATTTGATGACGACAAAATAATAGATTCAAAACAAACAACTCGGTGGCATATTGGGAG 358339
Qy 1100 GGATATCGAATGTTGAGATTGTGCACCTTCAAAGTTGCTTTTAAAGCCGACACCACTATCG 1159
Db 358340 GAATTAGCTGTGAATCCCAATAGTTTAAAGATTGCGATTAAGCCAAACCTTCAATAG 358399
Qy 1160 GGTGTAACAGAACACCGT---GTCAAGGAGCGTGCAGAACGTTGAGCTTCTAGCAAGAG 1216
Db 358400 GTAAAAGCAAAAACCAATAAATTTAAACCTTAGAAAATGTTGAAATTTGAAATTTGAAG 358459
Qy 1217 GCGCCATGACCATGCTGCTCGCCCTCGAGCTGCTCTGTTGGTGAATCCATGGCCGCT 1276
Db 358460 GAAGACAGCAGCAGTTATGTTTCAAGGATTGTTCCAGTGGCTGAAGCAATGGTTGCTA 358519
Qy 1277 TGTCTCTATGACACGAGCTGAT 1298
Db 358520 TAACCTTAGCTGATTGATGAT 358541

RESULT 15

US-09-064-693A-20
; Sequence 20, Application US/09064693A
; Patent No. 6210937
; GENERAL INFORMATION:
; APPLICANT: Ward, Thomas E.
; TITLE OF INVENTION: DEVELOPMENT OF GENETICALLY
; TITLE OF INVENTION: ENGINEERED BACTERIA FOR PRODUCTION
; TITLE OF INVENTION: OF A SPECIFIC PLASTICS PRECURSOR
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. Gary Goodson
; ADDRESSEE: INEL--Lockheed Martin Idaho
; ADDRESSEE: Technologies Co.
; STREET: P.O. Box 1625
; CITY: Idaho Falls
; STATE: Idaho
; COUNTRY: USA
; ZIP: 83415-3810
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: Toshiba Satellite Pro T2150CDS
; OPERATING SYSTEM: Windows95
; SOFTWARE: Word Perfect 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/064,693A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: W. Gary Goodson
; REGISTRATION NUMBER: 22,387
; REFERENCE/DOCKET NUMBER: Lit-PI-296
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (208)526-9469
; TELEFAX: (208)526-8339

; INFORMATION FOR SEQ ID NO: 20:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1239 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-09-064-693A-20

Query Match 10.5%; Score 170; DB 3; Length 1239;
Best Local Similarity 49.9%; Pred. No. 4.5e-44;

Matches 540; Conservative 0; Mismatches 505; Indels 37; Gaps 3;

Qy 233 CTGGAACACGTTTGGAAACTTCTTTCAGGTTGCAACCTATGTTGTAATCTCATGCGGGTG 292
Db 128 CTGGAACACCAATTTGGCAAACTCTTTTCGCTAAACCACTTCGGCGAATCGCACGCGCTG 187
Qy 293 GTGTTGGTGTGTTATCATGTTGTTCCACCTAGTAATTCCTACTCTCTAGGAGACCTTAC 352
Db 198 CGCTCGGCTGCATCGTCGATGTTTCCGCGAGGCAATCCGCTGACGGAAGCGGACCTGC 247
Qy 353 AAGTTGAATCGATCGAAGACGCGCCGACAGAGCAGAATAAACCCTCCACAAGGAAGAGA 412
Db 248 AACATGACCTCGACGCTGCTCGCTTGGGACATCGCGCTATACCAACCGCGCCGAGC 307
Qy 413 CTGATACATGCAAAATTTCTGTGAGGACACATGAAGGGGTGACTACTGTGAACGCCAATTC 472
Db 308 CGGATCAGGTCAAAATTTCTCTCGGTGTTTTTGAAGGCGTTACTACCGGCACCAAGCATTTG 367
Qy 473 TTGTTATTGTCCTCAAAACACAGATCAAAATAGCAGTGCATCACCGTGAATACCCATGTGT 532
Db 368 GTTGTGATGCAAAACACTGACCGGCTCTCAGGATTTACAGTGCATTAAGACGTTT 427
Qy 533 ACCGACCTTCTATGACGACGCAACTTATGACTTCAAGTACGCTGTGTAGAGCTGTACAGG 592
Db 428 TCCGTCCAGGCAATGCGATTTACACCTACGAACAAATACGGTCTCGCGGATTTATCGC 487
Qy 593 GAGTGGGAGTCTCTCGGCGAAGAAACCGTTGGAAGGGTGGTGCAGGGGCCCTCCCCA 652
Db 488 GCGTGGAGCGTCTTCCGCGCGGAAACCGCCATGCGCGTGGCGGACAGGACTATTGCCA 547
Qy 653 AGAAATTTCTTAAGCTCAAAATGTGATTTAGAGATCTTGTCTGTTTGTTCCTCAAGTGCATC 712
Db 548 AAAATATCTCGCGAATAATTTGTAATGAAATCCGTGGCTGCCTGACCCAGATGGCG 607
Qy 713 AGTTGTGCTCCCAAGAGACCGGTGATTTATGGGTCTGTAACCTTTGGAAACAGATAGAGA 772
Db 608 ACATTCCGCTGGATATCAAAAGAC-----TGTTCGAGGTTCAGC 646
Qy 773 GCAACATCGTTAGATGCTCTGATCCAGAGTACGAGAAAGATGATAGACGCAATCGACA 832
Db 647 AAATCCGTTTTTTTGGCGGACCCCGACAAAATCGACGCTTAGACGAGTTGATGCGTG 706
Qy 833 GAGTACAGTTTCGAGGGGATTCGGTTCGGTGGAGTATCATCGCTCGCTAGAAACGTTTC 892
Db 707 CGCTGAAAAAAGAGGGCGACTCCATCGGCGCTAAAGTCAACGTTGTTCCAGTGGCGTTC 766
Qy 893 CTGCGGGCTCGTTCCTCTGTGTTTCGAAAGCTCGAATCGAACTCGCAAAAGCTATGC 952
Db 767 CTGCGGAGCTTGGGAGCGCGTCTTTGACCGCTGGATGCTGACATCGCCCATCGCCTGA 826
Qy 953 TTTCTATTCTCGGAGCAACGGGTTTCGAGATTGGCAGCGGATTCGCGGGACCGCATTTGA 1012
Db 827 TGAGCATCAACGCGGTGAAAGCGTGGAAATTTGGCGACGCGCTTTGACGTTGGTGGCGCTGC 886
Qy 1013 CAGGAAGTGAAGCATAATGATGAGTTTTTATATGATAAGCTGGAAGTGTTCAGGACAGCA 1072
Db 887 GCGGACGCGAAGACCGCGATGAAATCAACCAAGA-----CGTTTCCAGA 931
Qy 1073 CTAATCGCTCGGTGTGTGCGAGGAGGATATCGAATGTTGAGATTTGCACTTCAAG 1132
Db 932 GCAACCATGCGGCGGCAATCTCGCGGTATCAGCAGCGGCGGAGCAATCATTTGCCATA 991
Qy 1133 TTGCTTTTAAAGCGCACACCATCTATCGGGGTGAACAGAACACCGTGTCAAGGAGCGTC 1192

Db	992	TGGCGCTGAAACCGACCTCCAGCATACCGTGCCGGGTCGTACCATTAACCGCTTTGGCG	1051
Qy	1193	AGAACGTTGAGCTTCTAGCAAGAGGGCGCCATGACCCATGCGTCGCCCTCGAGCTGTC	1252
Db	1052	AAGMAAGTTGAGATGATCACCAAGGCCGCTCAGATCCCTGTGTGGGATCCGCGCAGTGC	1111
Qy	1253	CTGTGGTGAATC-CATGGCCCGTTGCTCCTCATGGACCAAGCTGATGGCGCACGTGGCT	1311
Db	1112	CGATCGCAAGACGGAATCTGGCGATCGTTTTTAATGGATCACCTGTTACGGCAACGGGCG	1171
Qy	1312	CA 1313	
Db	1172	CA 1173	

Search completed: August 27, 2005, 04:27:18
Job time : 318 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2005, 00:49:04 ; Search time 1018 Seconds
(without alignments)
10450.947 Million cell updates/sec

Title: US-10-677-179-7

Perfect score: 1626

Sequence: 1 gcaagagctcagcttcgtct.....caaaaaaaaaaaaaaaaaaa 1626

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7331713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1626	100.0	1626	18	US-10-677-179-7
2	1385	85.2	1496	18	Sequence 7, Appli
3	1382.2	85.0	1541	20	Sequence 30604, A
4	1321.8	81.3	1845	20	Sequence 138335
5	872	53.6	1792	20	Sequence 138337
6	870.4	53.5	1635	18	US-10-425-115-171490
7	867.2	53.3	1603	18	US-10-677-179-1
					Sequence 1, Appli
					Sequence 32244, A

Sequence 99207, A
Sequence 11, Appl
Sequence 8982, Ap
Sequence 1811, Ap
Sequence 27, Appl
Sequence 5, Appli
Sequence 171487,
Sequence 13775, A
Sequence 9032, Ap
Sequence 13579, A
Sequence 43897, A
Sequence 42555, A
Sequence 26364, A
Sequence 1812, Ap
Sequence 9009, Ap
Sequence 8754, Ap
Sequence 46069, A
Sequence 44740, A
Sequence 1, Appli
Sequence 7, Appli
Sequence 9218, A
Sequence 45609, A
Sequence 34106, A
Sequence 9882, Ap
Sequence 12647, A
Sequence 6185, Ap
Sequence 1892, Ap
Sequence 1892, Ap
Sequence 31517, A
Sequence 41598, A
Sequence 41210, A
Sequence 24843, A
Sequence 42063, A
Sequence 6135, Ap
Sequence 47170, A
Sequence 20439, A
Sequence 5, Appli

ALIGNMENTS

RESULT 1

US-10-677-179-7
; Sequence 7, Application US/10677179
; Publication No. US20040082050A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Pember, Stephen O.
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-A
; CURRENT APPLICATION NUMBER: US/10/677,179
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: US/09/743,207
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: PCT/US99/16353
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 1626
; TYPE: DNA
; ORGANISM: Zea mays
US-10-677-179-7

Query Match 100.0%; Score 1626; DB 18; Length 1626;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCACGAGCTCAGCTTCGTCTCTCGCCGGCGGCGAGCAACTATCATCTTCATTAG 60
DB 1 GCACGAGCTCAGCTTCGTCTCTCGCCGGCGGCGAGCAACTATCATCTTCATTAG 60
QY 61 CTCATCCAACTCTATTCGATGACGACCGTGCAGCCAGCCACAGCGAGTGGCCACTCAGG 120
DB 61 CTCATCCAACTCTATTCGATGACGACCGTGCAGCCAGCCACAGCGAGTGGCCACTCAGG 120
QY 121 GCACGGCTCGACCCCGCGGATCGGCGCTTCGTGGAGTTTGGCCAGCCTCTCTCTCC 180
DB 121 GCACGGCTCGACCCCGCGGATCGGCGCTTCGTGGAGTTTGGCCAGCCTCTCTCTCC 180
QY 181 CTCGCTTCGCGCTGCACCGTGCAGCTCTGCTGCGCTTAGAGTGAAGGCACTCGGAAC 240
DB 181 CTCGCTTCGCGCTGCACCGTGCAGCTCTGCTGCGCTTAGAGTGAAGGCACTCGGAAC 240
QY 241 ACGTTTGGAAACTACTTTTCAGTTGCAACCTATGTTGAATCTCATGGGGTGGTGGT 300
DB 241 ACGTTTGGAAACTACTTTTCAGTTGCAACCTATGTTGAATCTCATGGGGTGGTGGT 300
QY 301 TGTGTTATCAGTGGTTCACCTAGATTCACACTCAGTGGAGGAGACTTCAAGTTGAA 360
DB 301 TGTGTTATCAGTGGTTCACCTAGATTCACACTCAGTGGAGGAGACTTCAAGTTGAA 360
QY 361 CTCGATCGAAGACGCGCGGACAGACAGATAAATCTCCACAGGAAGAGACTGATACA 420
DB 361 CTCGATCGAAGACGCGCGGACAGACAGATAAATCTCCACAGGAAGAGACTGATACA 420
QY 421 TGCATAATTCGTTCAGGAGACATGAAGGGTGAATCTGGAACGCCAATTTCTTTATT 480
DB 421 TGCATAATTCGTTCAGGAGACATGAAGGGTGAATCTGGAACGCCAATTTCTTTATT 480
QY 481 GTCCCAACACAGATCAATAGGCACTGATCACCGTGAATAGCCAAATGTCACCGCT 540
DB 481 GTCCCAACACAGATCAATAGGCACTGATCACCGTGAATAGCCAAATGTCACCGCT 540
QY 541 TCTCATGACAGACCAACTTATGACTTCAAGTACGGTGTAGAGCTGTACAGGAGTGGG 600
DB 541 TCTCATGACAGACCAACTTATGACTTCAAGTACGGTGTAGAGCTGTACAGGAGTGGG 600
QY 601 AGGTCCTCGGCGAGAAAAACCGTTGGAAGGGTGGCTGCAGGGGCCCTCCCAAGAAAT 660
DB 601 AGGTCCTCGGCGAGAAAAACCGTTGGAAGGGTGGCTGCAGGGGCCCTCCCAAGAAAT 660
QY 661 CTTAAGCTCAATATGGATATAGATCTTGTGTTGTTTCCAAAGTGCATCAGTTGTG 720
DB 661 CTTAAGCTCAATATGGATATAGATCTTGTGTTGTTTCCAAAGTGCATCAGTTGTG 720
QY 721 CTCCAGAGACGCGGTTGATTTATGGGTCTGTAACTTTTGGAAACAGATAGAGCAACATC 780
DB 721 CTCCAGAGACGCGGTTGATTTATGGGTCTGTAACTTTTGGAAACAGATAGAGCAACATC 780
QY 781 GTTAGATGTCCTGATCCAGAGTACGACAGAGATGATAGCGCAATCGACAGAGTACGA 840
DB 781 GTTAGATGTCCTGATCCAGAGTACGACAGAGATGATAGCGCAATCGACAGAGTACGA 840
QY 841 GTTCAGGGGATTCGGTGGTGGATGATCACATCGCTCGCTAGAAAGTTCTTCGGGG 900
DB 841 GTTCAGGGGATTCGGTGGTGGATGATCACATCGCTCGCTAGAAAGTTCTTCGGGG 900
QY 901 CTCGGTCTCTCTGTTTCGACAAAGCTCGAACTCCGAACTGGCAAAAGCTATGTTCTATT 960
DB 901 CTCGGTCTCTCTGTTTCGACAAAGCTCGAACTCCGAACTGGCAAAAGCTATGTTCTATT 960
QY 961 CTGCGAGCAACGGGTTCCAGATTCGACAGCGGATTCGCGGGACCGAATTCACAGGAAGT 1020
DB 961 CTGCGAGCAACGGGTTCCAGATTCGACAGCGGATTCGCGGGACCGAATTCACAGGAAGT 1020
QY 1021 GAGCATATGATGATGTTTATATGATTAAGCTGGAAGTGCAGACGCGACTAATCGC 1080
DB 1021 GAGCATATGATGATGTTTATATGATTAAGCTGGAAGTGCAGACGCGACTAATCGC 1080
QY 1081 TCGGGTGGTGTGCAGGGGGATATCGAATGTTGAGATTTGTGCACTTCAAAAGTTGCTTTT 1140

DB 1081 TCGGGTGGTGTGCAGGGGGATATCGAATGTTGAGATTTGCACTTCAAAAGTTGCTTTT 1140
QY 1141 AAGCCGACACATCTATCGGGGTGAACAGAACACCGTGTCAAGGGAGCGTCAAGACGTT 1200
DB 1141 AAGCCGACACATCTATCGGGGTGAACAGAACACCGTGTCAAGGGAGCGTCAAGACGTT 1200
QY 1201 GAGCTCTACGACAGGGCGCCATGACCCATGCGTGGTCCCTCGAGCTGTTCTGTGGTG 1260
DB 1201 GAGCTCTACGACAGGGCGCCATGACCCATGCGTGGTCCCTCGAGCTGTTCTGTGGTG 1260
QY 1261 GAATCCATGCGCGCTTGGTCTCATGACAGCTGATGGCGCACGCTGCTCAGTGGCGAG 1320
DB 1261 GAATCCATGCGCGCTTGGTCTCATGACAGCTGATGGCGCACGCTGCTCAGTGGCGAG 1320
QY 1321 ATGTTGCGGCTCAATATCTGCACTTCAAGAACCGAGTTGGCTCTTTCTAGCAGAGCAGC 1380
DB 1321 ATGTTGCGGCTCAATATCTGCACTTCAAGAACCGAGTTGGCTCTTTCTAGCAGAGCAGC 1380
QY 1381 ACACCTGATGAGCTCGCGCCAAATTTTATCATTTATCATAGTAATAAGTAGTCAAGCGT 1440
DB 1381 ACACCTGATGAGCTCGCGCCAAATTTTATCATTTATCATAGTAATAAGTAGTCAAGCGT 1440
QY 1441 GGCTTGGTTCCTTCTCTTGCAACCGTAGTTTGTGTTTTTTTCCCGCAAGTGTGATG 1500
DB 1441 GGCTTGGTTCCTTCTCTTGCAACCGTAGTTTGTGTTTTTTTCCCGCAAGTGTGATG 1500
QY 1501 CGATGAAGTGAATAAGGCACTTGGTTTCCTGTCGATTTGTACAGCTTTCATATAATGTA 1560
DB 1501 CGATGAAGTGAATAAGGCACTTGGTTTCCTGTCGATTTGTACAGCTTTCATATAATGTA 1560
QY 1561 TCTACTTCGAAGATGATGATTTTATAGATGTTGGCTTGTGAAGACAAAAA 1620
DB 1561 TCTACTTCGAAGATGATGATTTTATAGATGTTGGCTTGTGAAGACAAAAA 1620
QY 1621 AAAAAA 1626
DB 1621 AAAAAA 1626

RESULT 2

US-10-425-114-30604
; Sequence 30604, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 30604
; LENGTH: 1496
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73045G02_FLI
US-10-425-114-30604

Query Match 85.2%; Score 1385; DB 18; Length 1496;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 1490; Conservative 0; Mismatches 5; Indels 82; Gaps 2;
QY 12 GCTTCGCTCTCTCGCGCGCGGCAACTATCATCTTCATAGTCCATCCATC 71
DB 1 GCTTCGCTCTCTCTCGCGCGCGGCAACTATCATCTTCATAGTCCATCCATC 60

```
QY 72 TATTCCGATGACAGACCGTGCCCAAGCCACAGCAGAGTGGCGCACTCAACGGGACCGGCTCGC 131
DB 61 TATTCCGATGACAGACCGTGCCCAAGCCACAGCAGAGTGGCGCACTCAACGGGACCGGCTCGC 120
QY 132 ACCCGCGCGATCGCGGCTTGGTGGAGTTGGCCCGAGGCTCTCTCCCTCGGCTTGGC 191
DB 121 ACCCGCGCGATCGCGGCTTGGTGGAGTTGGCCCGAGGCTCTCTCCCTCGGCTTGGC 180
QY 192 CGTGACACCGCTCGCGCACTGCTCGCTAGAGTGAAGGCATCTGGAACACACGTTTGGAAA 251
DB 181 CGTGACACCGCTCGCGCACTGCTCGCTAGAGTGAAGGCATCTGGAACACACGTTTGGAAA 240
QY 252 CTACTTTTCAGGTGCAACCTATGTTGAATCTCATGCGGGTGGTGGTGGTGGTGGTGGTGGT 311
DB 241 CTACTTTTCAGGTGCAACCTATGTTGAATCTCATGCGGGTGGTGGTGGTGGTGGTGGTGGT 300
QY 312 TGGTTGTCCACCTAGAAATTCGCTCACTGAGGAGACCTACAGTTGAACTCGATCGAAG 371
DB 301 TGGTTGTCCACCTAGAAATTCGCTCACTGAGGAGACCTACAGTTGAACTCGATCGAAG 360
QY 372 ACCGCCCGGACAGAGCAGAAATTAACCTCCCAAGGAAGGAGACTGATACATGCAAAATTC 431
DB 361 ACCGCCCGGACAGAGCAGAAATTAACCTCCCAAGGAAGGAGACTGATACATGCAAAATTC 420
QY 432 GTGAGGACACATGAAGGGGTGACTACTGGAACGCCAATTCCTGTTATTTGTCCTCCAAACAC 491
DB 421 GTGAGGACACATGAAGGGGTGACTACTGGAACGCCAATTCCTGTTATTTGTCCTCCAAACAC 480
QY 492 AGATCAAAATAGGACGTGATCAACGTTGAATAGCCAAATGTGTACCGACCTTCTCATGCGAGA 551
DB 481 AGATCAAAATAGGACGTGATCAACGTTGAATAGCCAAATGTGTACCGACCTTCTCATGCGAGA 540
QY 552 CGCAACTTATGACTTCAAGTACGTTGTAGAGCTGTACAGGAGGTGGAGGTCTCGGG 611
DB 541 CGCAACTTATGACTTCAAGTACGTTGTAGAGCTGTACAGGAGGTGGAGGTCTCGGG 600
QY 612 CAGAAAAACCGTTGGAAGGGTGGCTGCGAGGGGCTCCCAAGAAAAATTTCTTAAGCTCAA 671
DB 601 CAGAAAAACCGTTGGAAGGGTGGCTGCGAGGGGCTCCCAAGAAAAATTTCTTAAGCTCAA 660
QY 672 ATGTGGATTAGAGATCTTGCTGTTGTTTTCAAAGTGCAATCAGTTGTGCTCCCAAGAGA 731
DB 661 ATGTGGATTAG-----AGATAGAGAGCAACATCGTTAGATGCC 671
QY 732 CGCGGTTGATTATGGTCTGTAACTTTGGAAACAGATAGAGCAACATCGTTAGATGCC 791
DB 672 -----AGATAGAGAGCAACATCGTTAGATGCC 699
QY 792 TGATCCAGAGTACGACAGAAAGATGATAGACGCAATCGACAGATACAGATTCGAGGGGA 851
DB 700 TGATCCAGAGTACGACAGAAAGATGATAGACGCAATCGACAGATACAGATTCGAGGGGA 759
QY 852 TTGCGTGGTGGAGTGATCACAATGCGTGGCTAGAAAACGTTCTCGCGGGCTCGGTTCTCC 911
DB 760 TTGCGTGGTGGAGTGATCACAATGCGTGGCTAGAAAACGTTCTCGCGGGCTCGGTTCTCC 819
QY 912 TGTGTTCCAGAGCTCGAATCGAATCGCAAGCTATGCTTTCTATTCTCGCGAGCA 971
DB 820 TGTGTTCCAGAGCTCGAATCGAATCGCAAGCTATGCTTTCTATTCTCGCGAGCA 879
QY 972 CGGGTTCCAGAGTGGCAGCGAATTCGCGGGACCGACTTGACAGAAAGTGAGCATAATGA 1031
DB 880 CGGGTTCCAGAGTGGCAGCGAATTCGCGGGACCGACTTGACAGAAAGTGAGCATAATGA 939
QY 1032 TGAGTTTATATGGAATAGGCTGGAAGTGTACAGACAGCACTAATCGCTCGGGTGGT 1091
DB 940 TGAGTTTATATGGAATAGGCTGGAAGTGTACAGACAGCACTAATCGCTCGGGTGGT 999
QY 1092 GCAGGGAGGATATCAATGTTGAGATGTGCACTTCAAGTGTCTTTTAAGCCGACACC 1151
DB 1000 GCAGGGAGGATATCAATGTTGAGATGTGCACTTCAAGTGTCTTTTAAGCCGACACC 1059
QY 1152 ATCTATCGGGGTGAAACAGAACACCGTGTCAAGGGAGCGTCAGAACGTTGAGCTTCTAGC 1211
```

```
DB 1060 ATCTATCGGGGTGAAACAGAACACCGTGTCAAGGGAGCGTCAGAACGTTGAGCTTCTAGC 1119
QY 1212 AAGAGGGCGGCATGACCCATGCGTCCCGCTCGAGCTGTTCTGTGGTGGAAATCCATGCG 1271
DB 1120 AAGAGGGCGGCATGACCCATGCGTCCCGCTCGAGCTGTTCTGTGGTGGAAATCCATGCG 1179
QY 1272 CGCGTTGGTCTCTCATGGACACAGCTGATGGCGCACGTGGCTCAGTCCGAGATGTTCCGCGCT 1331
DB 1180 CGCGTTGGTCTCTCATGGACACAGCTGATGGCGCACGTGGCTCAGTCCGAGATGTTCCGCGCT 1239
QY 1332 CAATACTGCATCTCAAGAACCAAGTTGGCTCTTTTCTAGCAGAGGACAGACACCTGATGA 1391
DB 1240 CAATACTGCATCTCAAGAACCAAGTTGGCTCTTTTCTAGCAGAGGACAGACACCTGATGA 1299
QY 1392 GCTCGGCGCAATTTTATCATATGATTAAGTAGCTCAAGCGTGGCTTGGTTG 1451
DB 1300 GCTCGGCGCAATTTTATCATATGATTAAGTAGCTCAAGCGTGGCTTGGTTG 1359
QY 1452 CTGTGCTCTGCGACCGTAGTTTGG-TTTTTTTTTCCCGCAAGTGTGATGCGATGAAGTG 1510
DB 1360 CTGTGCTCTGCGACCGTAGTTTGGTTTTTTTTTCCCGCAAGTGTGATGCGATGAAGTG 1419
QY 1511 AATAAGGCATCTGGTTTCTGTGCATTTGTACACGTTTTCATATATGTAATCTACTTCGA 1570
DB 1420 AATAAGGCATCTGGTTTCTGTGCATTTGTACACGTTTTCATATATGTAATCTACTTCGA 1479
QY 1571 AGATGATGATTTTAT 1587
DB 1480 AGATGATGATTTTAT 1496
```

RESULT 3

```
US-10-425-115-138335
; Sequence 138335, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 138335
; LENGTH: 1541
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MFT4577_57639C.1
US-10-425-115-138335
```

```
Query Match 85.0%; Score 1382.2; DB 20; Length 1541;
Best Local Similarity 93.2%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 28; Indels 82; Gaps 2;
QY 12 GCTTCGCTCTCTCTCGCGCGCGGCGGCACTATCATCTTCAATTAGCTCATCCAATC 71
DB 1 GCTTCGCTCTCTCTCGCGCGCGGCGGCACTATCATCTTCAATTAGCTCATCCAATC 60
QY 72 TATTCCGATGACAGACCGTGCCCAAGCCACAGAGTGGCGCACTCACGGGCAACGGCTCGC 131
DB 61 TATTCCGATGACAGACCGTGCCCAAGCCACAGAGTGGCGCACTCACGGGCAACGGCTCGC 120
QY 132 ACCCGCGCGATCGCGGCTTGGTGGAGTTGGCCCGAGCTTCTCCCTCCGCTCGGCTCGC 191
DB 121 ACCCGCGCGATCGCGGCTTGGTGGAGTTGGCCCGAGCTTCTCCCTCCGCTCGGCTCGC 180
QY 192 CGTGACACCGCTCGCGCACTGCTCGCTAGAGTGAAGGCATCTGGAACACAGTTTGGAAA 251
```

Db 181 CGTGCACCGCTGCCCACTGCTCGCTAGAGGTGAAGCATCTGGAACACGTTTGGAAA 240
Qy 252 CTACTTTTCAGGTTGCAACCTAATGTTGAATCTCATGCGGGGTGGTGTGTTGTATACAG 311
Db 241 CTACTTTTCAGGTTGCAACCTAATGTTGAATCTCATGCGGGGTGGTGTGTTGTATACAG 300
Qy 312 TGGTGTCCACTAGAAATTCACCTACACAGGCGAGACCTACAGTTGAACCTCGATCGAAG 371
Db 301 TGGTGTCCACTAGAAATTCACCTACACAGGCGAGACCTACAGTTGAACCTCGATCGAAG 360
Qy 372 ACGGCCCGGACAGACAGAGCAATAAACCCTCCACAGGAAGGAGACTGATACATCAAAAATTC 431
Db 361 ACGGCCCGGACAGACAGAGCAATAAACCCTCCACAGGAAGGAGACTGATACATCAAAAATTC 420
Qy 432 GTCAGGGAACATGAAGGGGTGAATCTGGAACCGCAATTCCTGTTATGTGTCCTCAACAC 491
Db 421 GTCAGGGAACATGAAGGGGTGAATCTGGAACCGCAATTCCTGTTATGTGTCCTCAACAC 480
Qy 492 AGATCAAAATAGCAGTGATCACCGTGAATAGCCAAATGTGTACCGACCTTCTCATGCAGA 551
Db 481 AGATCAAAATAGCAGTGATCACCGTGAATAGCCAAATGTGTACCGACCTTCTCATGCAGA 540
Qy 552 CGCAACTTATGACTTCAAGTAGCGGTGTAGAGCTGTACAGGAGGTGGGAGTCTCTCGG 611
Db 541 CGCAACTTATGACTTCAAGTAGCGGTGTAGAGCTGTACAGGAGGTGGGAGTCTCTCGG 600
Qy 612 CAGAAAAACCGTTGGAAGGGGTGCGAGGGGCCCTCCCAAGAAAAATCTTAAGCTCAA 671
Db 601 CAGAAAAACCGTTGGAAGGGGTGCGAGGGGCCCTCCCAAGAAAAATCTTAAGCTCAA 660
Qy 672 ATGTGGATTAGAGATCTTGTGCTTTGTTTCCAAAGTGCAATCAGGTGTGCTCCCAAGA 731
Db 661 ATGTGGATTAG----- 671
Qy 732 CGCGGTTGATTATGCGTCTGTAACTTTGGAAACAGATAGAGAGCAACATCGTTAGATGCC 791
Db 672 -----AGATAGAGAGCAACATCGTTAGATGCC 699
Qy 792 TGATCCAGAGTAGCAGAGAGATGATAGACCAATCGACAGGTACAGTTCCAGGGGA 851
Db 700 TGATCCAGAGTAGCAGAGAGATGATAGACCAATCGACAGGTACAGTTCCAGGGGA 759
Qy 852 TTCGCTCGGTGAGTGATCACATGCTCGCTAGAAAACGTTCTCGCGGGCTCGGTTCTCC 911
Db 760 TTCGCTCGGTGAGTGATCACATGCTCGCTAGAAAACGTTCTCGCGGGCTCGGTTCTCC 819
Qy 912 TGTGTTCCGAACAGCTCGAATCCGAATCGCAAAAGCTATGCTTTCTATTCCTCGAGCAA 971
Db 820 TGTGTTCCGAACAGCTCGAATCCGAATCGCAAAAGCTATGCTTTCTATTCCTCGAGCAA 879
Qy 972 CGGCTTCGAGATTGGCAGCGGATTCGCGGGACCGACTTGACAGAAAGTGAGCATATGA 1031
Db 880 CGGCTTCGAGATTGGCAGCGGATTCGCGGGACCGACTTGACAGAAAGTGAGCATATGA 939
Qy 1032 TGAGTTTATATGGAATAGGCTGGAAGTGTGAGGACACGCACTAATCGCTCGGGTGGT 1091
Db 940 TGAGTTTATATGGAATAGGCTGGAAGTGTGAGGACACGCACTAATCGCTCGGGTGGT 999
Qy 1092 GCAGGAGGGATATCGAATGTTGAGATTGTGCACTTCAAAAGTTGCTTTAAAGCCGACACC 1151
Db 1000 GCAGGAGGGATATCGAATGTTGAGATTGTGCACTTCAAAAGTTGCTTTAAAGCCGACACC 1059
Qy 1152 ATCTATCGGGTGAACAGAACACCGTGTCAAGGAGCGTCAAGACGTTGAGTCTTAGC 1211
Db 1060 ATCTATCGGGTGAACAGAACACCGTGTCAAGGAGCGTCAAGACGTTGAGTCTTAGC 1119
Qy 1212 AAGAGGGCCCATGACCCATCGCTCGCCCTCGAGCTTTCCTGTGGTGGAAATCCATGGC 1271
Db 1120 AAGAGGGCCCATGACCCATCGCTCGCCCTCGAGCTTTCCTGTGGTGGAAATCCATGGC 1179
Qy 1272 CGCGTTGGTCTCATGGACACAGCTGATGGCGCACGTTGGCTCAGTGGAGATGTTCCGCT 1331
Db 1180 CGCGTTGGTCTCATGGACACAGCTGATGGCGCACGTTGGCCCGAGATGTTCCGCT 1239

Qy 1332 CAATACTGCATTTCAAGAACCACTTGGCTCTTTCTAGCAGAGCAGACACACTGATGA 1391
Db 1240 CAATGCTGCATTTCAAGAACCACTTGGCTCTTTCTAGCAGAGCAGACACACTGATGA 1299
Qy 1392 GCTCGGCCCAAAATTTTATCATTTATCATAGTAATAAGTAGCTCAAGCGTGGCTTGGTTG 1451
Db 1300 GCTCGGCCCAAAATTTTATCATTTATCATAGTAATAAGTAGCTCAAGCGTGGCTTGGTTG 1359
Qy 1452 CTTGTCTCTTGCACCGTAGTTTTTGTGTTTTTTT-TTCCGCAAGTGTGATCGGATGAAGTG 1510
Db 1360 CTTGTCTCTTGCACCGTAGTTTTTGTGTTTTTTTCCCGCAAGTGTGATCGGATGAAGTG 1419
Qy 1511 AATAAGGCACTTGTGTTCCCTGTCATTTGTACAGTTTTCATATAATCTAATCTACTTCGA 1570
Db 1420 AATAAGGCACTTGTGTTCCCTGTCATTTGTACAGTTTTCATATAATCTAATCTACTTCGA 1479
Qy 1571 AGATGATGCATTTTATAGATGTGGCTTGTGAAAGACAAAAAAAAGAAAAA 1621
Db 1480 AGATGATGCATTTTATAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1530

RESULT 4

US-10-425-115-138337
; Sequence 138337, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 138337
; LENGTH: 1845
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_57640C.1
US-10-425-115-138337

Query Match 81.3%; Score 1321.8; DB 20; Length 1845;

Best Local Similarity 87.3%; Pred. No. 0;
Matches 1574; Conservative 0; Mismatches 17; Indels 212; Gaps 3;

Qy 11 AGCTTCGTCTCTCTCGCGCGCGGCGGAGGCAATCATCATCTTCTAGCTCATCAAT 70
Db 43 AGCTTCGTCTCTCTCGCGCGCGGCGGAGGCAATCATCATCTTCTAGCTCATCAAT 102
Qy 71 CTATTCCGATGACACCGTGCACCAAGCCACAGCAGGTGGCGCACTCAAGGCAAGGCTCG 130
Db 103 CTATTCCGATGACACCGTGCACCAAGCCACAGCAGGTGGCGCACTCAAGGCAAGGCTCG 162
Qy 131 CACCCCGCGGATCGGCGCTTGTGGAGTTTCCCGAGCTCTCTCCCTCCGCTTCG 190
Db 163 CACCCCGCGGATCGGCGCTTGTGGAGTTTCCCGAGCTCTCTCCCTCCGCTTCG 222
Qy 191 CCGTGACCGCTGCGGCACTGCTCGCTTAGAGTGAAAGGCACTCTGGAAACACGTTTGGAA 250
Db 223 CCGTGACCGCTGCGGCACTGCTCGCTTAGAGTGAAAGGCACTCTGGAAACACGTTTGGAA 282
Qy 251 ACTACTTTCAGGTTGCAACCTATCGTGAATCTCAATGGGGGTGGTGTGTTGTGTATCA 310
Db 283 ACTACTTTCAGGTTGCAACCTATCGTGAATCTCAATGGGGGTGGTGTGTTGTGTATCA 342
Qy 311 GTGGTTGTCACCTAGAAATTCCTCACTGAGGCGAGCACTACAAAGTTGAACCTCGATCGAA 370
Db 343 GTGGTTGTCACCTAGAAATTCCTCACTGAGGCGAGCACTACAAAGTTGAACCTCGATCGAA 402

```
QY 371 GAAGCGCCGACAGAGCAGAGTAACCTCCCAAGGAAGGAGTGTATCATCATCAAAATTC 430
|
Db 403 GAAGCCCGGTTCAGAGCAAGTAACCTCCCAAGGAAGGAGTGTATCATCATCAAAATTC 462
|
QY 431 TGTCAAGGACACATGAAGGGGTGACTACTGGAACGCAATCTTGTATTTGTCCTCAACA 490
|
Db 463 TGTCAAGGACACATGAAGGGGTGACTACTGGAACGCAATCTTGTATTTGTCCTCAACA 522
|
QY 491 CAGATCAAAATAGCGAGTATCACCGTGAATAGCCAAATGTGTACCGACCTTCTCATGAG 550
|
Db 523 CAGATCAAAATAGCGAGTATCACCGTGAATAGCCAAATGTGTACCGACCTTCTCATGAG 582
|
QY 551 ACAGCACTTATGACTTCAAGTACGGGTGTAGAGCTGTACAGGGAGGTGGAGGTCTCGG 610
|
Db 583 ACAGCACTTATGACTTCAAGTACGGGTGTAGAGCTGTACAGGGAGGTGGAGGTCTCGG 642
|
QY 611 GCAGAAAAACCGTTGGAAGGGGTGCTGAGGGGCCCTCCCAAGAAAAATCTTAAAGCTCA 670
|
Db 643 GCAGAAAAACCGTTGGAAGGGGTGCTGAGGGGCCCTCCCAAGAAAAATCTTAAAGCTCA 702
|
QY 671 AATGTGGATTAGAGATCTTGTCTGTTGTTTCCAAAGTGCATCAGGTGTGCTCCAGAG 730
|
Db 703 AATGTGGATTAGAGATCTTGTCTGTTGTTTCCAAAGTGCATCAGGTGTGCTCCAGAG 762
|
QY 731 ACAGCGTTGATTATGGGTCTGTAACTTTGGAAACAGATAGAGCAACATCGTTAGATGTC 790
|
Db 763 ACAGCGTTGATTATGGGTCTGTAACTTTGGAAACATATAGAGAGCAACATCGTTAGATGTC 822
|
QY 791 CTGATCCAGAGTACCGAGAGAGATGATAGACCAATCGACAGATACGAGTTCGAGGG 850
|
Db 823 CTGATCCAGAGTACCGAGAGAGATGATAGACCAATCGACAGATACGAGTTCGAGGG 882
|
QY 851 ATTTCGGTGGTGGAGTATCATCGGTCTGTAGAAAGTTCTCGGGGTCTGGTCTC 910
|
Db 883 ATTTCGGTGGTGGAGTATCATCGGTCTGTAGAAAGTTCTCGGGGTCTGGTCTC 942
|
QY 911 CTGATCTCGAAGCTCGAATCCGAATCGGAAAGCTATGTTCTATTTCTCGGAGCA 970
|
Db 943 CTGATCTCGAAGCTCGAATCCGAATCGGAAAGCTATGTTCTATTTCTCGGAGCA 1002
|
QY 971 ACAGGTTCCAGATTCGCGAGGATTCGCGGAGCCGACTTGACAGAAAGTGAAGCAATG 1030
|
Db 1003 ACAGGTTCCAGATTCGCGAGGATTCGCGGAGCCGACTTGACAGAAAGTGAAGCAATG 1062
|
QY 1031 ATGATGTTTATATGATAGGCTGCAAGTGTGAGCAGCAGGACTAATCGCTCGGTGG-- 1088
|
Db 1063 ATGATGTTTATATGATAGGCTGCAAGTGTGAGCAGCAGGACTAATCGCTCGGTGG 1122
|
QY 1089 ----- 1088
|
Db 1123 TCAGGTTTTCCTTCATCCTTAATAAACCCCTATCGATTGTGACTGTGGCATTTGTT 1182
|
QY 1089 -----TGTGAGGAGGAGATATCGAATGTTG 1114
|
Db 1183 CTCATAAAAAAATGGAATCGGTTTGTTCGTAATGTGAGGAGGAGATATCGAATGTTG 1242
|
QY 1115 AGATTGTGCACTTCAAGTGTCTTTTAAGCCGACACCATCTATCGGGGTGAAACAGACA 1174
|
Db 1243 AGATTGTGCACTTCAAGTGTCTTTTAAGCCGACACCATCTATCGGGGTGAAACAGACA 1302
|
QY 1175 CCGTGTCAAGGAGCGTCAAGACGTTGAGCTTCTAGCAAGAGGGGGCCATGACCCATGCG 1234
|
Db 1303 CTGTGTCAAGGAGCGTCAAGACGTTGAGCTTCTGTGCAAGAGGGGGCCATGACCCATGCG 1362
|
QY 1235 TCGCCCTCG----- 1244
|
Db 1363 TCGCCCTCGAGGTAAATGCAACCGTGAATAGCGCTCCTCGTGTAGTGTGCTGAACAT 1422
|
QY 1245 ----- 1244
|
Db 1423 TCTTTTAAATTTCTTCCCGGAGCTGCGACTGCGATACAGAGTTTGCTCTTGAAGTGGCT 1482
|
QY 1245 -----AGCTGTTCTGTGTGGAATCCATGCGCGGTTGTTCTCATGAGCAGCTGATG 1299
```

```
Db 1483 TCTGAGCTGTTCTGTGTGGAATCCATGCGCGGTTGTCCTCGTGGACCAAGCTGATG 1542
|
QY 1300 GCGCACGTGGTCTAGTGTGCGAGATGTTCCGCTCAATACTGCACTTCAAGAACCAAGTTGGC 1359
|
Db 1543 GCGCACGTGGCCAGTGTGCGAGATGTTCCGCTCAATACTGCACTTCAAGAACCAAGTTGGC 1602
|
QY 1360 TCTTCTAGCAGAGCGCAGACACCTGATGAGCTCGCGCCAAATTTTATCATTTATCAT 1419
|
Db 1603 TCTTCTAGCAGAGCGCAGACACCTGATGAGCTCGCGCCAAATTTTATCATTTATCAT 1662
|
QY 1420 AGTAATAGTAGCTCAAGCGTGGCTTGTGTTGCTTGTCTCTTGCACCGTAGTTTGTGTTT 1479
|
Db 1663 AGTAATAGTAGCTCAAGCGTGGCTTGTGTTGCTTGTCTCTTGCACCGTAGTTTGTGTTT 1722
|
QY 1480 TTTT-TTCCCGCAAGTGTGATGCGATGAAGTGAATAAGGCACTTGGTTTCTGTGCATTT 1538
|
Db 1723 TTTTCCCGCCCAAGTGTGATGCGATGAAGTGAATAAGGCACTTGGTTTCTGTGCATTT 1782
|
QY 1539 GTACACGTTTTCATATAATGTAATCTACTTCCGAAGATGATGATTTTATAGATGTGGCTT 1598
|
Db 1783 GTACACGTTTTCATATAATGTAATCTACTTCCGAAGATGATGATTTTATAGATGTGGCTT 1842
|
QY 1599 GTG 1601
|
Db 1843 GTG 1845
```

RESULT 5

```
US-10-425-115-171490
; Sequence 171490, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 171490
; LENGTH: 1792
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_87981C.1
US-10-425-115-171490
```

```
Query Match 53.6%; Score 872; DB 20; Length 1792;
Best Local Similarity 80.6%; Pred. No. 3,1e-261;
Matches 1032; Conservative 0; Mismatches 245; Indels 3; Gaps 1;
```

```
QY 83 CGACGTCGCCCAAGCAGCAGGTCGGGCACTCACGGGACGGCTCGCACCCCGCGGA 142
|
Db 173 GCGCCGTGTGCGACCGCCCGGTGTCCGCCAGGCGTCCACACCGTTCTCCCCCGGGA 232
|
QY 143 TCGGCGCTTGTCTGAGTGTTCGCCCGACCTCTCTCCCTCCGCTTGGCCGTGACCCGCT 202
|
Db 233 TAGGCGGCTCCCGAGTCCGCCC---CCAGTCCCTCGGTATCCGTGCGCGCGGTC 289
|
QY 203 GCGGACGTCTGCTGAGTGAAGGATCTGGAACACAGCTTTGGAAACTACTTTTCAAG 262
|
Db 290 GCGGCGCGCAGCCTTAGAGGTGAAGCATCGGGAATGTGTCGGGAATCTACTTCCAGG 349
|
QY 263 TTGCAACTATGTTGAATCTCATGCGGGTGTGTTGTTGTTATCATGTTGTTGTCAC 322
|
Db 350 TTGCAACTATGTTGGAATCTCATGAGGGGTGTGTTGTTGTTATCATGTTGTTGTCAC 409
|
QY 323 CTAGAAATTCACCTCACTGAGGACACCTTACAAGTTGAATCGATCGAAGACGCGCCGAC 382
|
```

Db 410 CCAGAAATTCCTCTCACTGAGGAGACATGCAAGTAGAACTCGATAGAAAGAGCTCCGGGTC 469
 Qy 393 AGAGCAGAAATTAACCTTCCACAGAGAGGAGAGCTGATACATGCAAAATTTCTGTACAGGACAC 442
 Db 470 AAAGTAGAAATTAACAACCCCAAGAGAGAGAGCTGATACATGCAAAATTTCTATCAGGAGAC 529
 Qy 443 ATGAAGGGTGACTACTGAGAGCGCAATTCCTGTTTATTTGTCCTCCAAACACAGATCAAAATAG 502
 Db 530 ATGATGGGATGACTACTGATACCAATTTACGCTTTTGTCCCAACACAGATCAAAAGGG 589
 Qy 503 GCAGTGATCAGCGTGAATAGCAATGTGTACCGACCTTCTCATGCGAGACCAACTTATG 562
 Db 590 GTGGTGATTAAGTGAATGTCTAAGGGGTACAGACCATCCATGCGAGATGCAACCTATG 649
 Qy 563 ACTTCAAGTACGGTGTAGAGCTGTACAGGAGAGTGGAGGTCTCTCGGCGAGAAAAACCG 622
 Db 650 ACTTCAAGTATGAGCTGTGACAGGAGGTGGAAGGTCTATCAGCCAGAGAAACCA 709
 Qy 623 TTGGAAGGGTGGCTGACGGGCGCTCCCAAGAAATTTCTTAAGCTCAAAATGTGGATAG 682
 Db 710 TTGGCAGGGTGGCTGACGAGCTCTTGGCAAGAAATTTCTAAGCTCAAAATGTGGATAG 769
 Qy 683 AGATCTTGTCTGTTGTTTCCAAAGTGCATCAGGTTGTCTCCCAAGAGACGCGGTTGATT 742
 Db 770 AGATCTTGGCATTTGTTTCTAAGGTGCACCAAGTCTTCTCCAGAGATGCAGTTGATT 829
 Qy 743 ATGGGTCTGTAACTTTGGAACAGATAGAGAGCAACATCGTTAGATGTCCTGATCCAGAGT 802
 Db 830 ATGAGACTGTAACTTTGGAACATATAGAGAGCAACATCGTTAGATGTCCTGATCCAGAA 889
 Qy 803 ACAGCAGAAAGATGATAGACGCAATCGACAGAGTACGAGTTTCGAGGGGATTTCCGTCGGTG 862
 Db 890 ATGCAGAGAGATGATCTCGCATTTGATACGTTACGATGTTAGAGAGATTTCAATTTGGTG 949
 Qy 863 GAGTGATCACATGCGTCTGAGAAAGCTTCTCGCGGGCTCGGTTCTCTGTGTTTCGACA 922
 Db 950 GGGTCGTACATGCAATTCGAAGAAATGTTCTCGTGGTCTTGGCTCTCTGTTTGGACA 1009
 Qy 923 AGCTCGAATCCGAATCGCAAGCTGATGCTTTCTATTCCTCGAGGCAACGGGTTTCGAGA 982
 Db 1010 AACTTGAAGCTGAATCGCAAGCAAGCATGCTTTCTCTTCTGCAAGCAAGGGGTTTGAGA 1069
 Qy 983 TTGCGAGCGGATTCGCGGGGACCGACTTGCAGAGAGTGAGCATAAATCATGAGTTTATATA 1042
 Db 1070 TTGGCAGTGGGTTCTGCTGACGAGCTTTACGGAAGTGAGCATAAATGATGAGTTCTATA 1129
 Qy 1043 TGGATAAGGCTGGAAAGTGTTCAGGACACGGAATAATCGCTCGGGTGTGTGAGGAGGGA 1102
 Db 1130 TGGATGAGGCTGGAAATGTGAGGACACGAACTAATCGCTCAGGCGGTGTTTCAGGGAGGGA 1189
 Qy 1103 TATCGAATGTTGAGATTTGTGCACTTTCAAGTTGCTTTTAAAGCGGACACCATCTATCGGG 1162
 Db 1190 TATCAATAGGTGAATATTTACTTTCAAGTGGCTTTTAAAGCAACAGCAACTATCGGAA 1249
 Qy 1163 TGAAACAGAACACCGTGTCAAGGAGCGTCAGAACGTTTGAGCTTCTAGCAAGAGGGGCGCC 1222
 Db 1250 AGAACCAAAATACTGTGTCAAGGAGCATGAGATGTTGAACTTTTGCAGAGGGGCGGCC 1309
 Qy 1223 ATGACCCATGCGTCCCGCTCGAGCTGTTCCTGTGGTGGAAATCCATGCGCGTGTGGTCC 1282
 Db 1310 ATGACCCCTGTGTTGCTCCTCGAGCTGTTCTATGTTGGAAATCCATGCGTGTGGTCC 1369
 Qy 1283 TCATGGACAGCTGATGGGCGACGTGGCTCAGTGGAGATGTTCCGCTCAATACTGACAC 1342
 Db 1370 TGATGGACAGCTATGCGCATATTTGCCCGAGTGTGAGATGTTTCCGCTGAAACCTTGCCC 1429
 Qy 1343 TCCAGAACCACTTGGCTCT 1362
 Db 1430 TACAAGGCCCATTTGGCTCT 1449

; Sequence 1, Application US/10677179
 ; Publication No. US20040082050A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cahoon, Rebecca E.
 ; APPLICANT: Falco, Saverio Carl
 ; APPLICANT: Pember, Stephen O.
 ; TITLE OF INVENTION: Choriplate Biosynthesis Enzymes
 ; FILE REFERENCE: BB-1159-A
 ; CURRENT APPLICATION NUMBER: US/10/677,179
 ; CURRENT FILING DATE: 2003-10-02
 ; PRIOR APPLICATION NUMBER: US/09/743,207
 ; PRIOR FILING DATE: 2001-01-04
 ; PRIOR APPLICATION NUMBER: PCT/US99/16353
 ; PRIOR FILING DATE: 1999-07-20
 ; PRIOR APPLICATION NUMBER: 60/093,611
 ; PRIOR FILING DATE: 1998-07-21
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 1
 ; LENGTH: 1635
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; US-10-677-179-1
 Query Match 53.5%; Score 870.4; DB 18; Length 1635;
 Best Local Similarity 80.5%; Pred. No. 9.2e-261;
 Matches 1031; Conservative 0; Mismatches 246; Indels 3; Gaps 1;
 Qy 83 CGACCGTGCACCAAGCCACAGCAGTGGCGCAGCTCACGGGACGCGCTCGCACCCCGCGCA 142
 Db 59 CGCCCGTGTGCGAGCGCGGTGTCCGCCAGGGGCTCCACACGCGTTTCTCCCGCGCGGA 118
 Qy 143 TCGCGCGCTTGTGGAGTTTGGCCCGAGCTCTCTCCCTCCCGCTTCCCGCTGACCGCT 202
 Db 119 TAGCGCGCTCCCGAGTCCGCCC---CCACGCTCCCTCCGCTTATCCGTCGCGCGCTC 175
 Qy 203 GCGCAGCTGTGCGCTAGAGTGAAGCATCTGGAACACGTTTGGAAACATCTACTTTCCAGG 262
 Db 176 GCGCGGCTCGAGCTAGAGTGAAGCATCTGGAACATCTACTTTCCAGG 235
 Qy 263 TTGCAACTATGTGAAATCTCATGGGGGTGGTGTGTTGTTGTTATCAGTGGTTGTCCAC 322
 Db 236 TTGCAACTATGTGGAATCCATGGAGGGGTGTGGTGTGTTATCAGTGGTGTCCAC 295
 Qy 323 CTAGAAATTCACCTCAGTGGAGCAGCTACAGTTGAATCTGATGAGAGCGGCGCCGAC 382
 Db 296 CCAGAAATTCCTCTCAGTGGAGCAGCATGCAAGTAGAACTCGATAGAAGAGCTCCGGTCC 355
 Qy 383 AGAGCAGAAATAACCTCCACAAAGAGAGACTGATACATGCAAAATTTCTGTACGGGACAC 442
 Db 356 AAGTAGAATTTACACCCCAAGAGAGAGACTGATACATGCAAAATTTCTATCAGGAGACAC 415
 Qy 443 ATGAAGGGGTGACTACTTGGAAACGCAATTTCTGTTATTTGTCCTCCAAACACAGATCAAAATAG 502
 Db 416 ATGATGGGATGACTACTTGGTACACCAATTCACGCTCTTTGTCCCAACACAGATCAAAAGG 475
 Qy 503 GCAGTGATCAGCTGAAATAGCCATGTGTACCGACCTTCTCATGCGAGAGCAACTTATG 562
 Db 476 GTGGTGATTAAGTGAATGTCTAAGCGGTACAGACCATCCCATGCGAGATGCAACCTATG 535
 Qy 563 ACTTCAAGTACGGTGTAGAGCTGTACAGGAGGTGGAGGTCTCCGGGACAGAAAAACCG 622
 Db 536 ACTTCAAGTATGAGTTAGAGCTGTGAGGAGGTGGAGGTCTATCAGCCAGAGAAACCA 595
 Qy 623 TTGGAAGGGTGGTGTGAGGGGCGCTCCCAAGAAAAATTTTAAAGCTCAAAATGTGATAG 682
 Db 596 TTGGCAGGGTGGCTGAGGAGCTCTTGCAGAAAAATTTTAAAGCTCAAAATGTGATAG 655
 Qy 683 AGATCTTGTGTTTGTTCAGAAAGTGCATCAGGTTGTGCTCCCAAGAGAGCGCGTTCATT 742
 Db 656 AGATCTTGGCATTTGTTTCTAAAGTGCACCAAGTCTGCTTCCAGAGATGAGCTTGAAT 715
 Qy 743 ATGGGTCTGTAACCTTTTGGAAACAGATAGAGAGCAACATCGTTAGATGTCCTGATCCAGAGT 802

1121 AGAAGCAAAATACGTGTCAAGGAGCATGAGGATGTTGAACCTTTTGGCAAGGGGGCGCC 1180
 1223 ATGACCCATCGGTGCGCCCTCGAGCTGTTCTGTGTGGTGAATTCATGAGCGCGTGTGTC 1282
 1181 ATGACCCCTGTGTGTGTCCTCGAGCTGTTCTATGTTGGAATCCATGCGTGTGTC 1240
 1283 TCATGGACCACTGATGCGGACGCTGCTCAGTCGAGATGTTGCGGCTCAATACTGCAC 1342
 1241 TGATGGACCACTCATGCGGATATTTGCCCACTGTGAGATGTTTCCGCTGAACCTTGCCC 1300
 1343 TTCAAGAACCACTGTTGGCTCT 1362
 1301 TACAAGACCCCATTTGGCTCT 1320

RESULT 8

US-10-437-963-99207/c
 ; Sequence 99207, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrew A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 99207
 ; LENGTH: 1353
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_9703C.1
 US-10-437-963-99207

Query Match 49.6%; Score 806.4; DB 19; Length 1353;
 Best Local Similarity 78.4%; Pred. No. 8.8e-241;
 Matches 1010; Conservative 0; Mismatches 236; Indels 42; Gaps 2;
 75 TCCGATGACGACCGTGCCCAAGCCACAGCAGGTGGCGCACTCAGCGGACGCGCTCGCAC 134
 1285 TCCGATGCGCGCCCAAGCTGTCGACGCGGTGGCGCGCTCTCCCGCGGGGG 1226
 135 CGCGCGATCGCGCGCTTGTGGAGTTTGGCCCGAGCTCTCTCCCTCGCGTTCGCGGT 194
 1225 CGCGCGGTTCGCGCGCTTCCGCGAGTTCGCGCGCGCT-----TCCCTCCGCTTCGCGT 1172
 195 GCACGCTGCGCGCACTGCTCGCTAGAGTGAAGCATCTGGAACACAGTTTGGAACTA 254
 1171 CGCGCGCGCGCGCGCTTCCGCTAGAGTGAAGCGTCTGCAATGATATTGGGAACCTA 1112
 255 CTTTCAGGTGCAACTATGATGTAATCTCATGGGGGTGGTGTGGTGTGTTATCAGTGG 314
 1111 CTTTCAGGTGCAACTATGATGAGAGTCTCATGGAGCGGTGTGGTTCGCTAATCAGTGG 1052
 315 TTGTCCACTAGAAATTCACCTCACTGAGGAGACCTACAAGTTGAATCGATCGAAGACG 374
 1051 ATGCCACCCAGAAATCCCACTTACTTGAAGCAGATATGCAAGTAGAACTCGACCGGAGACG 992
 375 GCCCGACAGACGACAATAAATCCACAGGAGGAGAGTATATCATATCAAAATTCCTGTC 434
 991 GCCAGCGACAGCAGAAATAACACCCCAAGAAAGAGAGTGCACCTTGCAAAATTCCTTC 932
 435 AGGGACACATGAAGGGGTGACTTACTTGGAAACGCGCAATCTTGTATTGTTCCCAACACAGA 494

931 AGGCACACATGAAGGAATGACCACCTGGGACACCAATTCATGTTTGTTCCTCCGAACACAGA 872
 495 TCAATAGGCACTGATCACCCTGAAATAGCAATGTGTACCGACCTTCTCATGCAAGCGC 554
 871 TCAGAGAGGGGGTGTATACAGTGAATGGCTAAGGCTTCAGACCTTTCATCATGAGATGC 812
 555 AACTTATGACTTCAAGTACCGTGTGTAGAGCTGTGTACAGGAGGTGGGAGGTCTCTCGGCAG 614
 811 AACTTATGACTTCAATACCGTGTGTAGAGCAGTGTGAGGAGGTGGAGAGATCATCAGCAAG 752
 615 AAAAACCGTTTGAAGGGTGGCTGCAGGGGCGCTCCCAAGAAAAATCTTAAAGCTCAAAATG 674
 751 AGAGACCATTTGAAGGGTGGCTGCAGGAGCTCTTGCAGAAAAATCTTAAAGCTCAAAATC 692
 675 TGGATTAGAGATCTTGTGCTGTTTGTTCCTCAAGTGCATCAGGTTGTGCTCCAGNAGACGC 734
 691 TGGAGTAGAGATCTTGGCGTGTGTGCTCAAGGTGCATCAAGTTGTACTACAGAGATGC 632
 735 GGTGATTATGGGTCTGTAACTTTTGGAAACAGATAGAGCAACATCGTTAGATGCTCCTGA 794
 631 CGTTGATTATGACACTGTAACTTGAACAGATAGAAAGCAACATTTGTTAGATGCTCCTGA 572
 795 TCCAGAGTACGACAGAGAGATGATAGACCAATCGACAGATGACAGTTCGAGGGGATC 854
 571 TCCAGAAATATGCAAGAGATGATGATGCAATCGATAAAGTACGAGTTAGAGGTGATTC 512
 855 GGTGCGTGGAGTGTACATGCTGCTAGAAAGTTCCTCGGGGCTCGGTTCTCCTGT 914
 511 GATTGGTGGTGGTGTGATGCAATTCGCAAGAAATGTTTCTCGTGGGATGGCTTCTCCTGT 452
 915 GTTCGACAACTCGAATCCGAATCGCAAGTGTGCTTCTTATTCCTCGGCAACCGG 974
 451 ATTTGACAACTTGAGGCTGAAATGGGCAAGTGTGCTTCTCTCTCGCAACGAGG 392
 975 GTTCGAGATTGGCAGCGGATTCGCGGACCGCACTTGCAGGAGTGCAGATATGATGA 1034
 391 GTTTGAGATCGGAGTGGATTTGCAGGTACTGACTACACTGGAAGTGCAGCATATGATGA 332
 1035 GTTTTATATGATAGGCTCGAAGTGTGCAGACACGCACTAATCGCTCGGGTGTGTGCA 1094
 331 GTTCTATATGATGAGGCTGGA-----A 308
 1095 GGAGGGATATCGAATGTTGAGATTTGCACATTCAGAAAGTTGCTTTTAAAGCGCACCATC 1154
 307 TGGAGGATATCAATGGTGAATATATATCTTCAAGTAGCTTTCAGGCCAACCGGAC 248
 1155 TATCGGGTGAACAGAACACCGTGTCAAGGGAGCGTGCAGACGTTGAGCTTCTAGCAAG 1214
 247 TATTGGGAAGAAACACATACTGTTTCAAGGGAGCATGAGGATGTTGAACTTTTAGCAAG 188
 1215 AGGGCGCATGACCCATGCGTCCCGCTCGAGCTGTTCTGTGTGTGGAATCCATGGCGGC 1274
 187 GGCGCGCACGACCATGTTGTCTCGCTCGCTGTCGAGTGTGAGTGCATGGCGGC 128
 1275 GTTGGTCTCATGACCACTGATGGCGCACGCTGGCTCAGTGCAGAGATGTTTCGCGCTCAA 1334
 127 ATTAGTCTCATGACCACTGATGGCGCATGTTGCTCAATGTGAGATGTTTCCACTGAA 68
 1335 TACTGCACTTCAAGAACCACTGCTCT 1362
 67 CTTGCTCTACAGAACCACTGCTCT 40

RESULT 9

US-10-677-179-11
 ; Sequence 11, Application US/106771179
 ; Publication No. US20040082050A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Caboon, Rebecca E.
 ; APPLICANT: Falco, Saverio Carl
 ; APPLICANT: Pember, Stephen O.
 ; TITLE OF INVENTION: Chorisemate Biosynthesis Enzymes
 ; FILE REFERENCE: BB-1159-A

164	QY	CCCCAGCCTCTCTCTCCCTCGGTTCCGCGGTGCACCGCTCCCGCACTGCTCGCCTAGAGG	223
90	Db	CCCTCCCCCGCGCTACTCCGACTCTCACTCCGTCCTCGTCTTCCCAAGAGACTTCACA	149
224	QY	TGAAGGCATCTGGAAAACAGGTTTGGAAACTACTTTTCAGGTTGCAAACTATGGTGAATCTC	283
150	Db	TACAGGCGGTGGAGTGACCTATGGAATCAGTTTCGTGTACACATATGGGAATCAC	209
284	QY	ATGCGGGTGGTGTGGTGTGTATTCAGTGGTTGTCCACTAGAAATTCACATCACTAGG	343
210	Db	ATGAGGAGGTGTGGTGTGTATTCATGATGTCCTCTCGCTTCCTCTCTCTGAAG	269
344	QY	CAGACCTACAGTTTGAATCGATCGAGACGCGCCGACAGACAGACGAATAACCTCCACAA	403
270	Db	CTGATATCAAGTGGATCTTGACAAAGGAGGCCAGGTACAGCCGAATTAACAACCTCTA	329
404	QY	GGAAGGACATGATACATGCAGAAATTCGTGAGGGACACATGAAGGGGTGACTACTCGAA	463
330	Db	GAAAGGACATGATACATGATAAATAATTTTCAGGAGTTTCCGAAGGAATCACTACTGGAA	389
464	QY	CGCCAAATCTTGTATTGTGCCAACAACAGATCAATATGGCAGTGATCACCGTGAATAAG	523
390	Db	CTCCAAATTCATGTATCTGTACCCCAATCTGATCAAGAGGACATGACTATACGAGATGG	449
524	QY	CCAATGTGTACCGACCTTCTCATGCAGACGCAACTTATGACTTCAAGTACGGTGTTAGAG	583
450	Db	CAGTAGCTTATAGGCCCTCCCATGCAGATGCTACCTATGACATCAAGTATGTGTCAAT	509
584	QY	CTGTACAGGAGGTGGGAGGTCTCTCGGCGACAAAAACCGTTGGAAAGGTGGCTGCAGGGG	643
510	Db	CAGTTCAGGGTGGTGGTAGATCTTCTGCAAGAGAAACAATTGGAAGGGTTCTTCTGTGG	569

QY 644 CCTCCCAAGAAATCTTAAGCTCAATGTGGATTAGAGATCTTGCGTTGTTTCCA 703
 DB 570 CTGTGCTAAGAAATCTTAAGAAATTTCTGGAACTGAGATTCTGGCCCTATGTCTC 629
 QY 704 AAGTCATCAGGTGTGCTCCAGAGACGGGGTTGATTATGGGTCTCTAACTTTGGAAC 763
 DB 630 AAGTTCAATGAATTTCTTCCAGAGGACCTTATTGATCATGACACTCTGACTCTTGATC 689
 QY 764 AGATAGAGAGCAATCGTTAGATGCTCTGATCCAGAGTAGCGAGAGAGATGATAGACG 823
 DB 690 AGATTGAGAGTAACATTTGTCGATGCCAGACCCCGAGTATGCGAGAGATGATATCTG 749
 QY 824 CAATCGACAGAGTACGATTCGAGGGGATTCGGTCGGTGGAGTGATCATCATGCTCGCTA 883
 DB 750 CAATTGATGCTGTGCGAGTGAGAGTGATTTCTGTTGGTGGTGTGTGACATGCAATGFGA 809
 QY 884 GAAACGTTCTCGCGGGTCTGCTCTCTGTTCTCGACAAGCTCGAATCCGAATCGGCAA 943
 DB 810 GGAACGTGTCACGAGTCTCGGTTCCACAGTATTGACAAACTTGAAGCTGAGCTGGCTA 869
 QY 944 AAGCTATGCTTTCTATTCCTGGAGCAACGGGTTCCAGATTGGCAGCGGATTCGCGGGA 1003
 DB 870 AAGCTGCAATGTCTATTCGCTCGCAACCAAGGGCTTTTCAGTTTGGTAGTGGGTTTCAGGCA 929
 QY 1004 CCGACTTCACAGGAAGTGCAGCATATGATGATGATTTTATATGATAAGGCTGGAAGTGCTA 1063
 DB 930 CTTTTTGTGCTGGAGTGGAACAATGATGATGATTTCTATATAGATGAACATGGAACACAA 989
 QY 1064 GGACACGGAATTAATCGCTCGGGTGTGTCAGGAGGGATATCGAATGTTGAGATTGTGC 1123
 DB 990 GAAACAAGAAATCGCTCTGGTGGATACAGGGTGGAAATTTCCAAATCGGGAAATCATTA 1049
 QY 1124 ACTTCAAGTTGCTTTTAAAGCGACACCAATCTATCGGGGTGAACAGAAACCGGTGCAA 1183
 DB 1050 ATATGAGAATAGCTTTTCAAGCCAAATCAACAAATTTGGAAGAAAGCAAAAGACTGTGACTC 1109
 QY 1184 GGGAGCGTCAGAACGTTGAGCTTCTAGCAAGAGGCGCCATGACCATGCGTCCGCCCTC 1243
 DB 1110 GAGATAAAGAAACAGAGTTTATAGCCGTGGTGGCCATGATCTTTGTGTTGTTGCCAA 1169
 QY 1244 GAGCTGTTCTCTGTGTTGAATCCATGGCGGGTGGTCTCTATGACACAGCTGATGGCG 1303
 DB 1170 GAGCTGTACCTATGTTAGAACCAATGTTAGCTTTGTTCTTGTGGACCAATTTGATGGCA 1229
 QY 1304 AGCTGGCTCAGTCCAGATGTTGCGGCTCAATCTGCACTTTCAGAACCAAGTTG 1357
 DB 1230 AATATGCGCAGTGAATCTTTTCCCGTAAACTCAGATTGTCAGAACCCCTTGG 1283

RESULT 11
 US-10-424-599-1811
 ; Sequence 1811. Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhen Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 1811
 ; LENGTH: 2094
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MFT3847_101634C.1
 US-10-424-599-1811
 Query Match 35.4%; Score 576.4; DB 18; Length 2094;

Best Local Similarity 67.7%; Pred. No. 1e-168;
 Matches 808; Conservative 0; Mismatches 386; Indels 0; Gaps 0;
 QY 164 CCCAGCCTCTCTCTCCCTCCGCTCGCGTGCACCGCTCCGCACTGCTCGCCTAGAGG 223
 DB 275 CCCTCTCCCCGGCTACTCCGACTCTCACTCCGCTCTCGCTTCCCAAGAGACTTCGCA 334
 QY 224 TGAAGGATCTGGAACACGTTTGGAACTACTTTAGGTTGCACCTATGTTGTAATCTC 283
 DB 335 TACAGCGCGTGGAGTACCTATGGAATCACTTCGTGTTTACAACTATATGGGAATCAC 394
 QY 284 ATGGGGTGTGTTGGTGTGTTATCAGTGTGTCACCTAGAAATTCACCTACTCAGG 343
 DB 395 ATGGAGAGGTTGTTGGTGTGTTATGATGATGTCCTCTCGCTTCTCTCTCTGAG 454
 QY 344 CAGACCTTACAGTTGAACTCGATCGAAGACGGCCGACAGAGCAGAGCAATAACCTCCCAA 403
 DB 455 CTGATATGCAAGTGGATCTTGAAGAGAGGAGGCGAGGTCAAGCCGAATTTACAACCTCTA 514
 QY 404 GGAAGGATCTGATACATGTCGAAATTTCTGTGAGGACACATGAGGGGTGACTACTCGAA 463
 DB 515 GAAAGGAGACTGATACATGTAATGTAATATTTTCAGAGTTTCCGAAGGAATCACTACTCGAA 574
 QY 464 CGCCAAATCTTGTATTGTCCCAAAACACAGATCAATAGGACAGTATCACCGTGAATAG 523
 DB 575 CTCCATTCATGATCTGTACCCCACTACTGATCAAGAGGACATGACTATAGCGAGATGG 634
 QY 524 CCAATGTGTACCGACCTTCTATGACGCAACTTATGACTTCAAGTACGGTGTGTAGAG 583
 DB 635 CAGTAGCTTATAGGCCCTCCATGTCAGATGCTACTATGATCAATGAGTATGTTGTAGAT 694
 QY 584 CTGTACAGGAGGTGGAGGTCTCCGGGCAAAAACCGTTGGAAGGTTGGTGCAGGG 643
 DB 695 CAGTTACAGGTTGGTGGTGTGATCTTCTGCAAGAGAAACAATTTGGAAGGTTGCTTCTGGTG 754
 QY 644 CCCTCCCAAGAAATTTTAAAGCTCAATGTGATTAGAGATCTTCTGCTGTTGTTTCCA 703
 DB 755 CTGTTGTAGAAATTCCTTAAGAAATTTCTGGAACCTGAGATTCGCGCTATGCTCTC 814
 QY 704 AAGTCATCAGTTGTGCTCCCAAGAGACGGGTTGATGATGTTGTTGTTGTTGTTGTTG 763
 DB 815 AAGTTCAAGATGTTTCTTCCAGAGGACCTTATGATCATGACACTCTGACTCTTGATC 874
 QY 764 AGATAGAGCACAATCGTTAGATGCTCTGATCCAGAGTACGAGAGCAAGATGATAGACG 823
 DB 875 AGATTGAGTAAATTTGTTGATGTCAGACCCGGAGTATGCAAGAGATGATATCTG 934
 QY 824 CAATCGACAGATACGAGTTTCGAGGGATTCGGTCCGTGAGTATGATCAGATGCTCGCTA 883
 DB 935 CAATTGATGCTGTGCGAGTGAGAGTGATTTCTGTTGGTGTGTTGTTGATGATGATGTA 994
 QY 884 GAAAGTTCTTCGCGGGCTCGGTTCTCTGTGTTGCAAGAGTCAAGTCCGAATCGGCA 943
 DB 995 GGAACCTGCCACGAGGTCTCGGTTCCACAGTATTGACAAACTTGAAGCTGAGCTGGCTA 1054
 QY 944 AAGCTATGCTTTCTATTCTCCGAGCAACGGTTCGAGATTGGCAGCGGATTCGCCGGGA 1003
 DB 1055 AAGCTCAATGTCATTCTCCCTGCAACCAAGGCTTTCAGTTTGGTGGTGGTTCGAGGCA 1114
 QY 1004 CCGACTTTCAGAGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1063
 DB 1115 CTTTTTGTGCTGGAGTGGAACACATGATGATGATGATGATGATGATGATGATGATGATGAT 1174
 QY 1064 GGACACGGAATATGCTCGGGTGTGTCAGGAGGGATATCAATGTTGAGATTGTGC 1123
 DB 1175 GAAACAAGAAACAAATCGCTCTGGTGGATACAGGGTGAATTTTCCAATGGGGAATCATTA 1234
 QY 1124 ACTTCAAGTTGCTTTTAAAGCGACACCACTATCGGGGTGAACAGACCAACCTGTCAC 1183
 DB 1235 ATATGAGAATAGCTTTTCAAGCCAACTCAACAAATTTGGAAGAAAGAAAGACTGTGACTC 1294
 QY 1184 GGGAGCGTCAGAACGTTTGTGAGCTTCTAGCAAGAGGCGCCCATGACCCATGCGTCCGCCCTC 1243


```
;
; LENGTH: 1015
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-677-179-5

Query Match      30.8%; Score 500.6; DB 18; Length 1015;
Best Local Similarity 80.3%; Pred. No. 3.6e-145;
Matches 587; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

Qy 631 GTGGCTGAGGGGCCCTCCCAAGAAAATCTTAAGCTCAAAATGTGGATTAGAGATCTTG 690
Db |||||
Qy 5 GAGGCTGAGGAGCTGTTGCAAGAAAATCTTAAGCTCAAAATGTGGATTAGAGATCTTA 64
Db |||||
Qy 691 TCGTTTGTTCACAAAGTCATCAGCTGTGCTCCAGAGACGGGTTGATTATGGGTCT 750
Db |||||
Qy 65 GATTTGTTTCCAAAGTCATCAGCTGTGCTCCAGAGACGGGTTGATTATGAAACT 124
Db |||||
Qy 751 GTAACTTTGGAAACAGATAGAGCAACATCTGTAGATGCTCTGATCCAGAGTAGCCAGAG 810
Db |||||
Qy 125 CTTACCTGGATCAGATAGAGCAACATTTGTAGATGCTCTGATCCAGATATATGCACAG 184
Db |||||
Qy 811 AAGATGATAGACCAATCGACAGAGTACGAGTTGAGGGGATTCGGTGGGTGAGTATC 870
Db |||||
Qy 185 AAGATGATGATGCAATGATAAAGTACGAGTTAATGGGAATTCGATTTGGTGGGTGGTC 244
Db |||||
Qy 871 ACATCGCTGCTAGAAAGCTTCTCGCGGCTCGGTTCTCTGTTTCGCAAGCTCGAA 930
Db |||||
Qy 245 ACATGCAATGCGCAAAATGTTCTCGTGGGCTTGGCTCTCTGTTATTTGACAAACTTGA 304
Db |||||
Qy 931 TCCGAATCGCAAAAGCTATGCTTCTATTTCTCGAGCAACGGGTTTCGAGATTGGCAGC 990
Db |||||
Qy 305 GCTCTACTGGCAAGGCTATGCTTCTTCTCGCAAGCAAGGGGTTTGAGATCGGTAGT 364
Db |||||
Qy 991 GAATTCGCGGAGCCGACTTGACAGGAAGTGAGCAATAATGATGATTTTATATGATTAAG 1050
Db |||||
Qy 365 GGAATTCGAGGTACTGACCTAACTGGAAGTGAGCAATAACGATGAGTTCTATATGACGAG 424
Db |||||
Qy 1051 GCTGAAGTGTACAGACAGGACTAATCGCTCGGGTGTGTGCGAGGAGGATATCGAAT 1110
Db |||||
Qy 425 GCTGGAATGTAAAGAACAGAACCAATCGCTCGGGGCTGTACAGGAGGAGATCAAAAT 484
Db |||||
Qy 1111 GTTGAGATTGTCACTTCAAAAGTTGCTTTTAAAGCCGACACCATCTATCGGGGTGAAAACAG 1170
Db |||||
Qy 485 GGTGAACCTATATACTTCAAGTAGCTTTCAGCCCAACAGCAACTATTTGGGNAAGCAA 544
Db |||||
Qy 1171 AACACCGTGTCAAGGAGGCTCAGAACGTTGAGCTTTAGCAAGAGGCGGCCATGACCCA 1230
Db |||||
Qy 545 AATACTGTAAACAAGGGATCATGAGGATATCGAACTTCTGACAAAGGGGTGCGCCATGACCCA 604
Db |||||
Qy 1231 TGCCTCGCCCTCGAGCTGTTCTCTGTTGGAATCCATGCGCGGTTGCTCTCATGGAC 1290
Db |||||
Qy 605 TGTGTGCTCCCTCGGGCTGTTCCAAATGGTGGAGACGATGGCTGCAATGGTCTCATGGAC 664
Db |||||
Qy 1291 CAGCTGATGGCGCAGCTGGCTCAGTGCAGAGATGTTTCGGGCTCAATPATCTGCACCTTCAAGAA 1350
Db |||||
Qy 665 CAGCTGATGGCACATGTTGCTCAGTGCAGAGATGTTCCCGCTCAACCTCGCCCTACAGAA 724
Db |||||
Qy 1351 CCAATTGGGCTC 1361
Db |||||
Qy 725 CCAATCGGCTC 735
Db |||||

RESULT 14
US-10-425-115-171487
; Sequence 171487, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants

; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 171487
; LENGTH: 724
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(724)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_87979C.1
US-10-425-115-171487

Query Match      26.2%; Score 425.6; DB 20; Length 724;
Best Local Similarity 79.1%; Pred. No. 8.6e-122;
Matches 542; Conservative 0; Mismatches 137; Indels 6; Gaps 3;

Qy 168 AGCTCTCTCTCCTCCCTCGCTTCGCGTGCAACCGCTGCGGCACCTGCTCGCCTAGAGGTGAA 227
Db |||||
Qy 41 ACCCTCAACCAACGAGGCAACCAACCTTCTCCGTCGCGCGCTCCAGCCTAGAGGTGAA 100
Db |||||
Qy 228 GGCACTCTGGAAACACAGCTTTTCGAAACTACTTTTCAGGTTGCAACCTATGCTGAATCTCATGG 287
Db |||||
Qy 101 GGCATCAGGAAATGTGTTTCGGAACTACTTTCAGGTTGCAACCTATGCGGAAATCCCATGG 160
Db |||||
Qy 288 GGGTGTGTGTTGTTTATCAGTGGTTGTTCACCTAGAAATTCACCTCATCTAGGAGGAGA 347
Db |||||
Qy 161 AGGGGTGTGTTGTTGCTTATCAGTGGCTGCCCAACAGAAATTCCTCTCACTGAGGAGGAGA 220
Db |||||
Qy 348 CCTACAAGTTGAATCGATCGAAGACGCGCCGAGACAGACAGATAACTCCACCAAGAA 407
Db |||||
Qy 221 CATGCAAGTAGAATCTCGATAGAGACGTCGCGGTCGCAAGTAGAATTAACAACCCCAAGAA 280
Db |||||
Qy 408 GGAGACTGATACATGCAAAATTTCTGTCAGGACACATGAAAGGGTGTACTCTGGAACGCC 467
Db |||||
Qy 281 GGAGACTGATACATGCAAAATTTCTGTCAGGACACATGATGGGATGACTACTGATACACC 340
Db |||||
Qy 468 AATTCTTTGTTATGTTCCAAACACAGATCAAAATAGGACGATGATCACCGTGAATATAGCAA 527
Db |||||
Qy 341 AATTCACTGTTTGTCCCAACACAGATCAAAAGGGTGTGATTAACAGTGAATGTCTAA 400
Db |||||
Qy 528 TGTGTACCGACCTTCTCATGACAGACCACTTATGACTTCAAGTACGCTGTGTAGAGCTGT 587
Db |||||
Qy 401 GGGGTACAGACCACTCCCATGACATGCAACTTATGACTTTTAAGTATGGAGTTAGAGCTGT 460
Db |||||
Qy 588 AC-AGGAGAGTGGGAGGTCTCTCGGCGAGAAAAACCCGTTGGAAGGGTGGCTGCGAGGGGCC 646
Db |||||
Qy 461 GCGNGNAGTGGNAGGTCTATCAGCAGAGAAACCATTTGCGAGGGTGGCTGCGAGAGCTC 520
Db |||||
Qy 647 TCCCAAGAAAATTTCTTAAGCTCAAAAT- - -GTGGATTAGAGATCTTTGTCTGTTGTTTCC 702
Db |||||
Qy 521 TTGCAAGAAAATTTCTTAAGCTCAAAATCANNGGAGTGGAGATCTNTTGGCATTTGTTTCT 580
Db |||||
Qy 703 AAGTGCATCAGTTGTGCTCCAGAGACGCGGTTGATTTATGCGTCTGTAACCTTTGGA 762
Db |||||
Qy 581 AAGTGCACCAAGTCGTAATTTCCAGAAAGTGCATGTTGATTTATGAGACTGTAAACCTTGAA 640
Db |||||
Qy 763 CAGATAGAGAGCAACATCGTTAGATGCTGATGATCCAGATACGCGAGAGAGATGATAGAC 822
Db |||||
Qy 641 CATATAGAGAGCAATCGTTAGATGCTGATGATCCCA- AATATGCGAGAGAGATGATGCT 699
Db |||||
Qy 823 GCAATCGACAGAGTACGAGTTCCGAG 847
Db |||||
Qy 700 GCCATTGATACGGTACGAGTTAGAG 724
Db |||||

RESULT 15
US-10-425-114-13775
; Sequence 13775, Application US/10425114
; Publication No. US2004003488A1
```

Search completed: August 27, 2005, 04:44:23
Job time : 1026 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 19:19:15 ; Search time 74 Seconds
(without alignments)
2242.165 Million cell updates/sec

Title: US-10-677-179-8

Perfect score: 2200

Sequence: 1 MTTVPKQQVAHSRARLAPR.....VAQCEMFALNTALQEPVGSF 429

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 385760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2200	100.0	429	3	AAV44893
2	1797.5	81.7	440	3	AAV44890
3	1541	70.0	431	8	ADK60818
4	1461	66.4	435	3	AG24018
5	1461	66.4	478	3	AG24017
6	1454	66.1	436	8	ADN73863
7	1454	66.1	436	8	ADT91427
8	1438.5	65.4	435	5	ABB91288
9	1257	57.1	307	3	AAV44895
10	1191.5	54.2	362	8	ADK60817
11	1191.5	54.2	362	8	ADN20024
12	1186.5	53.9	362	8	ADS29835
13	1182.5	53.8	362	8	ADS41780
14	1137	51.7	313	3	AG24019
15	1066	48.5	257	3	AAV44892
16	1035	47.0	360	8	ADS42623
17	981	44.6	356	8	ADS21386
18	938	42.6	376	8	ADK60821
19	938	42.6	376	8	ADS43492
20	934	42.5	378	8	ADP99010
21	930	42.3	432	8	ADK60819
22	929.5	42.2	412	8	ADR86076
23	885.5	40.4	358	6	ABU20907
24	886.5	40.3	473	8	ADN21319
25	862	39.2	361	6	ABU50323

26	860	39.1	361	4	AAU34639	Aau34639 E. coli c
27	860	39.1	361	6	ABU28699	Abu28699 Protein e
28	860	39.1	361	8	ADS45053	Ads45053 Bacterial
29	859	39.0	377	7	ABO66029	AbO66029 Klebsiell
30	858.5	39.0	357	8	ADO57720	AdO57720 Actinobac
31	858	39.0	352	6	ABU33103	Abu33103 Protein e
32	857	39.0	361	4	AAU36115	Aau36115 Klebsiell
33	855	38.9	361	6	ABU31873	Abu31873 Protein e
34	850	38.6	361	6	ABU41141	Abu41141 Protein e
35	850	38.6	362	7	ADF04007	AdF04007 Bacterial
36	846.5	38.5	357	6	ABU39054	Abu39054 Protein e
37	843	38.3	360	8	ADS42725	AdS42725 Bacterial
38	840.5	38.2	357	4	AAU35388	Aau35388 Haemophil
39	840.5	38.2	357	4	AAU35388	Aau35388 Haemophil
40	840.5	38.2	357	5	AAU31486	Aau31486 Haemophil
41	840.5	38.2	357	6	ABU30175	Abu30175 Protein e
42	840.5	38.2	357	8	ADK60820	AdK60820 Neurospor
43	840.5	38.2	359	8	ADS42183	AdS42183 Bacterial
44	840	38.2	361	6	ABU49470	Abu49470 Protein e
45	838	38.1	362	6	ABM70731	Abm70731 Photorhab

ALIGNMENTS

RESULT 1

AAV44893
ID AAV44893 standard; protein; 429 AA.

XX AAV44893;

DT 18-MAY-2000 (first entry)

DE Corn chorismate synthase encoded by clone csiln.pk0050.d11.

KW Corn cDNA clone csiln.pk0050.d11; chorismate synthase;

KW immunological screening; herbicide resistance; antibody; gene mapping;
KW corn.

OS Zea mays.

XX WO200005353-A2.

XX 03-FEB-2000.

XX 20-JUL-1999; 99WO-US016353.

XX 21-JUL-1998; 98US-0093611P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Cahoon RE, Falco SC, Pember SO;

XX WPI; 2000-182687/16.

XX N-PSDB; NAAZ50438.

XX New chorismate synthase polypeptides used to alter the level of the enzyme and thus the level of aromatic to non aromatic amino acids in transformed plants.

XX Claim 6; Page 33-34; 39pp; English.

XX The present sequence is corn chorismate synthase encoded by cDNA clone designated csiln.pk0050.d11. Manipulating either the amount or activity of this enzyme would affect the ratio of aromatic to non-aromatic amino acids in plants. Chimeric gene comprising the cDNA operably linked to regulatory sequences is used to transform host cells to alter the level of expression of chorismate synthase. The gene and its products may be used for immunological screening of cDNA expression libraries and to create transgenic plants which may also be herbicide resistant. Synthetic peptides derived from the gene are to raise antibodies, and used in screening assays to identify inhibitors which may be useful as herbicides

```
SQ Sequence 429 AA;
Query Match 100.0%; Score 2200; DB 3; Length 429;
Best Local Similarity 100.0%; Pred. No. 3.3e-206;
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTTVPKQVQVHRSRLAPRAIGALLEPAPASSSLRPAVHRCRTARLEVKASGNTFGNYF 60
DB 1 MTTVPKQVQVHRSRLAPRAIGALLEPAPASSSLRPAVHRCRTARLEVKASGNTFGNYF 60
QY 61 QVATYGESHGCGVGVISGCPPIPLTEADLQVELDRRRPQOSRITSTRKETDCKILSG 120
DB 61 QVATYGESHGCGVGVISGCPPIPLTEADLQVELDRRRPQOSRITSTRKETDCKILSG 120
QY 121 THEGVTGTPILVIVPNTDQIGSDHREITANYRPSHADATYDFKYGVRVAVGGGRSSGRK 180
DB 121 THEGVTGTPILVIVPNTDQIGSDHREITANYRPSHADATYDFKYGVRVAVGGGRSSGRK 180
QY 181 TVGRVAAGALPKKILKLCGLEILSFVSKVHVVLPEDADVGVSTLEQIESNIVRCPPDP 240
DB 181 TVGRVAAGALPKKILKLCGLEILSFVSKVHVVLPEDADVGVSTLEQIESNIVRCPPDP 240
QY 241 EYAEKMDAIDRVRVGRDVGCVITCVARNVPRGLSPVFDKLESELAKAMLSIPASNGF 300
DB 241 EYAEKMDAIDRVRVGRDVGCVITCVARNVPRGLSPVFDKLESELAKAMLSIPASNGF 300
QY 301 EIGSGFAGTDLTGSEHNDEFYMDKAGSVTRTNRSRGVGGISNVEIVHFKVAFKPTPSI 360
DB 301 EIGSGFAGTDLTGSEHNDEFYMDKAGSVTRTNRSRGVGGISNVEIVHFKVAFKPTPSI 360
QY 361 GVKQNTVSRERONVELLARGHDPVAPRAVPVSVESMAALVMDQLMAHVAQCEMPALNT 420
DB 361 GVKQNTVSRERONVELLARGHDPVAPRAVPVSVESMAALVMDQLMAHVAQCEMPALNT 420
QY 421 ALQEPVGSF 429
DB 421 ALQEPVGSF 429

RESULT 2
AA44890
ID AA44890 standard; protein; 440 AA.
AC AA44890;
XX
DT 18-MAY-2000 (first entry)
DE Corn chorismate synthase encoded by clone Chpc24.pk0002.hl.fis.
KW Corn cDNA clone chpc24.pk0002.hl.fis; chorismate synthase;
KW immunological screening; herbicide resistance; antibody; gene mapping;
KW corn.
OS Zea mays.
XX
PN WO200005353-A2.
XX
PD 03-FEB-2000.
XX
PF 20-JUL-1999; 99WO-US016353.
XX
PR 21-JUL-1998; 98US-0093611P.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI Cahoon RE, Falco SC, Pember SO;
XX
DR WPI; 2000-182687/16.
DR N-PSDB; AAZ50435.
XX
PT New chorismate synthase polypeptides used to alter the level of the
PT enzyme and thus the level of aromatic to non aromatic amino acids in
PT transformed plants.
```

```
XX Claim 6; Page 29-30; 39pp; English.
PS
XX The present sequence is corn chorismate synthase encoded by cDNA clone
CC designated chpc24.pk0002.hl.fis. Manipulating either the amount or
CC activity of this enzyme would affect the ratio of aromatic to non-
CC aromatic amino acids in plants. Chimeric gene comprising the cDNA
CC operably linked to regulatory sequences is used to transform host cells
CC to alter the level of expression of chorismate synthase. The gene and its
CC products may be used for immunological screening of cDNA expression
CC libraries and to create transgenic plants which may also be herbicide
CC resistant. Synthetic peptides derived from the gene are to raise
CC antibodies, and used in screening assays to identify inhibitors which may
CC be useful as herbicides
XX
SQ Sequence 440 AA;
Query Match 81.7%; Score 1797.5; DB 3; Length 440;
Best Local Similarity 81.9%; Pred. No. 9e-167;
Matches 349; Conservative 32; Mismatches 42; Indels 3; Gaps 2;
QY 4 VPKPQVQVHRSRLAPRAIGALLEPAPASSSLRPAV-HRCRTARLEVKASGNTFGNYFQV 62
DB 5 VSQPPVVSARASTRFLPRGIGALPESAP--TSLRLSVGRRRRRASSLEVKASGNVFGNYFQV 62
QY 63 ATYGESHGCGVGVISGCPPIPLTEADLQVELDRRRPQOSRITSTRKETDCKILSGTH 122
DB 63 ATYGESHGCGVGVISGCPPIPLTEADLQVELDRRRPQOSRITSTRKETDCKILSGTH 122
QY 123 EGVTTGTPILVIVPNTDQIGSDHREITANYRPSHADATYDFKYGVRVAVGGGRSSGRKT 182
DB 123 DGMVTGTPIHVFVNTDQRGDYSEMSKAYRPSHADATYDFKYGVRVAVGGGRSSARETI 182
QY 183 GRVAAGALPKKILKLCGLEILSFVSKVHVVLPEDADVGVSTLEQIESNIVRCPPDEY 242
DB 183 GRVAAGALAKKILKSGVEILAFVSKVHVVLPEDADVGVSTLEHIESNIVRCPPDEY 242
QY 243 AEKMDAIDRVRVGRDVGCVITCVARNVPRGLSPVFDKLESELAKAMLSIPASNGFEI 302
DB 243 AEKMDAIDTVRVGRDVGCVITCVARNVPRGLSPVFDKLESELAKAMLSIPASKGFEI 302
QY 303 GSGFAGTDLTGSEHNDEFYMDKAGSVTRTNRSRGVGGISNVEIVHFKVAFKPTPSIGV 362
DB 303 GSGFAGTDLTGSEHNDEFYMDKAGSVTRTNRSRGVGGISNVEIVHFKVAFKPTATIGK 362
QY 363 KQNTVSRERONVELLARGHDPVAPRAVPVSVESMAALVMDQLMAHVAQCEMPALNTAL 422
DB 363 KQNTVSRERONVELLARGHDPVAPRAVPVSVESMAALVMDQLMAHVAQCEMPALNTAL 422
QY 423 QEPVGS 428
DB 423 QEPVGS 428

RESULT 3
ADK60818
ID ADK60818 standard; protein; 431 AA.
XX
AC ADK60818;
XX
DT 20-MAY-2004 (first entry)
XX
DE Saccharomyces cerevisiae chorismate synthase.
XX
KW Apicomplexan parasite; malaria; toxoplasmosis; cryptosporidiosis;
KW eimeriosis; veterinary disease; chorismate synthase.
XX
OS Saccharomyces cerevisiae.
XX
FN US6699654-B1.
XX
PD 02-MAR-2004.
XX
```


PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139889P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142309P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.

PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 66.4%; Score 1461; DB 3; Length 435;
Best Local Similarity 67.5%; Pred. No. 8.1e-134;
Matches 276; Conservative 59; Mismatches 62; Indels 12; Gaps 2;

Qy 30 PASSSLRFAVHR-----CRT---ARLEVKASGNTFGNYFOVATYGESHGGVGCVI 77

Db 16 PGSSSLPSELRLSSPAVQISLRTQTRNFQIQATGSSYGTHFRVTFGSHGGVGCII 75

QY 78 SGCPRPIPTADQLQVELDRRRPGQSRITSTRKETDTCKILSGTHEGVTGTPILVIVPN 137
Db 76 DGCPPIPTESDQLDLDRRRPGQSRITTPRKETDTCTRISSGVSEGMTTGPPIHFVFN 135
QY 138 TQIGSDHREIANVRPSPHADATYDKYGVRAVQGGRRSGRKTGVRVAAGALPKKILKL 197
Db 136 TDQGLDYSEMVAVRPSPHADATYDMKYGVRSVQGGRRSARETIGRVAPGALAKKILKQ 195
QY 198 KCGLEILSPVSKVHVLPEDAVDYGVSVTLQEIENIVRCPEVAEKMIDAI DRVVRG 257
Db 196 FAGTEILAVSQVHHVLPVELLVDHENLTLEQIENNVRCNPPEVAEKMAIDAIVRTG 255
QY 258 DSVGVYITCAVRNVRGSLGSPVFDKLESELAKAMLSIPASNGFEICSGFAGTDLTGSEHN 317
Db 256 NSVGGVWTCIVRNAPRGLGTPVFDKLEAELAKACHSLPATKGFEGSGFAGTFLTGLEHN 315
QY 318 DEFYMDKAGSVTRTRNRSGVGGGINSVEIVHFKVAFKTPPSIGVKQNTVSRERQNVLL 377
Db 316 DEFYTDENGRIRTRNRSGGIQGGISNGEIIINRVAFKPTSTIGRKQNTVTRDKVETEMI 375
QY 378 ARGHRDPCVAPRAVPVWESMAALVMDQLMAHVAOCFALNTALQEPV 426
Db 376 ARGHRDPCVAPRAVPVWESMAALVMDQLMAHVAOCFALNTALQEPV 424

RESULT 5
AAG24017
ID AAG24017 standard; protein; 478 AA.
XX AC AAG24017;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 27534.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0128845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.

14-MAY-1999; 99US-0134370P.
18-MAY-1999; 99US-0134768P.
19-MAY-1999; 99US-0134941P.
20-MAY-1999; 99US-0135124P.
21-MAY-1999; 99US-0135353P.
24-MAY-1999; 99US-0135629P.
25-MAY-1999; 99US-0136021P.
27-MAY-1999; 99US-0136392P.
28-MAY-1999; 99US-0136782P.
01-JUN-1999; 99US-0137222P.
03-JUN-1999; 99US-0137528P.
04-JUN-1999; 99US-0137502P.
07-JUN-1999; 99US-0137724P.
08-JUN-1999; 99US-0138094P.
10-JUN-1999; 99US-0138540P.
10-JUN-1999; 99US-0138847P.
14-JUN-1999; 99US-0139119P.
16-JUN-1999; 99US-0139452P.
16-JUN-1999; 99US-0139453P.
17-JUN-1999; 99US-0139492P.
18-JUN-1999; 99US-0139454P.
18-JUN-1999; 99US-0139455P.
18-JUN-1999; 99US-0139456P.
18-JUN-1999; 99US-0139457P.
18-JUN-1999; 99US-0139458P.
18-JUN-1999; 99US-0139459P.
18-JUN-1999; 99US-0139460P.
18-JUN-1999; 99US-0139461P.
18-JUN-1999; 99US-0139462P.
18-JUN-1999; 99US-0139463P.
18-JUN-1999; 99US-0139750P.
18-JUN-1999; 99US-0139763P.
21-JUN-1999; 99US-0139817P.
22-JUN-1999; 99US-0139859P.
23-JUN-1999; 99US-0140353P.
24-JUN-1999; 99US-0140354P.
24-JUN-1999; 99US-0140695P.
28-JUN-1999; 99US-0140823P.
29-JUN-1999; 99US-0140931P.
30-JUN-1999; 99US-0141287P.
01-JUL-1999; 99US-0141842P.
01-JUL-1999; 99US-0142154P.
02-JUL-1999; 99US-0142055P.
06-JUL-1999; 99US-0142390P.
08-JUL-1999; 99US-0142803P.
09-JUL-1999; 99US-0142920P.
12-JUL-1999; 99US-0142977P.
13-JUL-1999; 99US-0143542P.
14-JUL-1999; 99US-0143624P.
15-JUL-1999; 99US-0144005P.
16-JUL-1999; 99US-0144085P.
16-JUL-1999; 99US-0144086P.
19-JUL-1999; 99US-0144325P.
19-JUL-1999; 99US-0144331P.
19-JUL-1999; 99US-0144332P.
19-JUL-1999; 99US-0144333P.
19-JUL-1999; 99US-0144334P.
19-JUL-1999; 99US-0144335P.
20-JUL-1999; 99US-0144352P.
20-JUL-1999; 99US-0144632P.
20-JUL-1999; 99US-0144884P.
21-JUL-1999; 99US-0144814P.
21-JUL-1999; 99US-0145086P.
21-JUL-1999; 99US-0145088P.
22-JUL-1999; 99US-0145085P.
22-JUL-1999; 99US-0145087P.
22-JUL-1999; 99US-0145089P.
22-JUL-1999; 99US-0145192P.
23-JUL-1999; 99US-0145145P.
23-JUL-1999; 99US-0145218P.
23-JUL-1999; 99US-0145224P.
26-JUL-1999; 99US-0145276P.
27-JUL-1999; 99US-0145913P.

PR 27-JUL-1999;	99US-0145918P.
PR 27-JUL-1999;	99US-0145919P.
PR 28-JUL-1999;	99US-0145955P.
PR 02-AUG-1999;	99US-0146386P.
PR 02-AUG-1999;	99US-0146388P.
PR 02-AUG-1999;	99US-0146389P.
PR 03-AUG-1999;	99US-0147038P.
PR 04-AUG-1999;	99US-0147204P.
PR 04-AUG-1999;	99US-0147302P.
PR 05-AUG-1999;	99US-0147192P.
PR 05-AUG-1999;	99US-0147260P.
PR 06-AUG-1999;	99US-0147303P.
PR 06-AUG-1999;	99US-0147416P.
PR 09-AUG-1999;	99US-0147493P.
PR 09-AUG-1999;	99US-0147935P.
PR 10-AUG-1999;	99US-0148171P.
PR 11-AUG-1999;	99US-0148319P.
PR 12-AUG-1999;	99US-0148341P.
PR 13-AUG-1999;	99US-0148565P.
PR 13-AUG-1999;	99US-0148684P.
PR 16-AUG-1999;	99US-0149368P.
PR 17-AUG-1999;	99US-0149175P.
PR 18-AUG-1999;	99US-0149426P.
PR 20-AUG-1999;	99US-0149722P.
PR 20-AUG-1999;	99US-0149723P.
PR 20-AUG-1999;	99US-0149920P.
PR 23-AUG-1999;	99US-0149930P.
PR 25-AUG-1999;	99US-0150566P.
PR 26-AUG-1999;	99US-0150884P.
PR 27-AUG-1999;	99US-0151065P.
PR 27-AUG-1999;	99US-0151066P.
PR 27-AUG-1999;	99US-0151080P.
PR 30-AUG-1999;	99US-0151303P.
PR 31-AUG-1999;	99US-0151438P.
PR 01-SEP-1999;	99US-0151930P.
PR 07-SEP-1999;	99US-0152363P.
PR 10-SEP-1999;	99US-0153070P.
PR 13-SEP-1999;	99US-0153758P.
PR 15-SEP-1999;	99US-0154018P.
PR 16-SEP-1999;	99US-0154039P.
PR 20-SEP-1999;	99US-0154779P.
PR 22-SEP-1999;	99US-0155139P.
PR 23-SEP-1999;	99US-0155486P.
PR 24-SEP-1999;	99US-0155658P.
PR 28-SEP-1999;	99US-0156458P.
PR 29-SEP-1999;	99US-0156596P.
PR 04-OCT-1999;	99US-0157117P.
PR 05-OCT-1999;	99US-0157753P.
PR 06-OCT-1999;	99US-0157865P.
PR 07-OCT-1999;	99US-0158023P.
PR 08-OCT-1999;	99US-0158232P.
PR 12-OCT-1999;	99US-0158369P.
PR 13-OCT-1999;	99US-0159293P.
PR 13-OCT-1999;	99US-0159294P.
PR 13-OCT-1999;	99US-0159295P.
PR 14-OCT-1999;	99US-0159323P.
PR 14-OCT-1999;	99US-0159330P.
PR 14-OCT-1999;	99US-0159331P.
PR 14-OCT-1999;	99US-0159637P.
PR 14-OCT-1999;	99US-0159638P.
PR 18-OCT-1999;	99US-0159584P.
PR 21-OCT-1999;	99US-0160741P.
PR 21-OCT-1999;	99US-0160767P.
PR 21-OCT-1999;	99US-0160768P.
PR 21-OCT-1999;	99US-0160770P.
PR 21-OCT-1999;	99US-0160814P.
PR 21-OCT-1999;	99US-0160815P.
PR 22-OCT-1999;	99US-0160980P.
PR 22-OCT-1999;	99US-0160981P.
PR 22-OCT-1999;	99US-0160989P.
PR 25-OCT-1999;	99US-0161404P.
PR 25-OCT-1999;	99US-0161405P.

PR 25-OCT-1999;	99US-0161406P.
PR 26-OCT-1999;	99US-0161359P.
PR 26-OCT-1999;	99US-0161360P.
PR 26-OCT-1999;	99US-0161361P.
PR 28-OCT-1999;	99US-0161920P.
PR 28-OCT-1999;	99US-0161922P.
PR 28-OCT-1999;	99US-0161993P.
PR 29-OCT-1999;	99US-0162142P.

Query Match

Best Local Similarity

Mismatches

Conservative

Score

DB

Length

66.4%;

67.5%;

59;

59;

1461;

3;

478;

276;

276;

59;

59;

12;

12;

2;

2;

Qy

30

PASSSLRPAVHR-----CRT--ARLEVASGNTFGNYFQVATYVGSHHGGVGCVI

77

Db

59

FGSSSLPSELRLSSPAVQISLRTQKQNFQIQATSSSYGTHFRVSTFGSHGGVGCII

118

Qy

78

SGCPPRIPLTEADLQVELDRRRPGQSRITSTRKETDTCKILSGTHEGVTGTPIIVPN

137

Db

119

DGCPPIPLTESDLQFDLDRRRPGQSRITTPRKETDTCRISSGVSEGWTGTPHVFVN

178

Qy

138

TDQIGSDHREIANVYRPSHADATYDFKYGVRVAVQGGRRSSGRKTVGRVAAGALPKILKL

197

Db

179

TDQKGLDYSEMSVAYRPSHADATYDMKYGVRVQGGRRSARETIGRVAPCALAKILKQ

238

Qy

198

KCGLLEILSFYSKVHVVLPEDAVDYGSVTLQIESNIVRCPDEYAEKMDAIDRVRVG

257

Db

239

PAGTEILAYYSQVHHVVLPELVDHENLTLEQIENNIVRCNPEYAEKMDAIDAVRTKG

298

Qy

258

DSVGGVITCVARNVPRGLGSPVFDKLESELAKAMLSIPASNGPEIGSGFAGTDITGSEHN

317

Db

299

NSVGGVVTICVRNAPRGLGTPVFDKLEAELAKACMSLPATKGFEGSGFAGTFLTGLEHN

358

Qy

318

DEFYMDKAGSVRTRTNRSQVQGGISNVEIVHFVKVAFKPTPSIGVKONTVSRERQNVELL

377

Db

359

DEFYTDENGRIRTRTNRSQGIQGGISNGEINNMVAFKPTSTIGRKQNTVTRDKVETEMI

418

Qy

378

ARGRHDPVAPRAVPVYESMAALVMDOLMAHVAQCEMFALNTALQEPV

426

Db

419

ARGRHDPVVPRAVPVWVESMVALVLDQLMAQYAQCHLFFPINPELQEPV

467

RESULT 6

ADN73863

ID

ADN73863

standard; protein; 436 AA.

XX

ADN73863;

DT

15-JUL-2004

(first entry)

DE

Thale cress protein repressed in E2Fa/Dpa expressing plants SeqID 1758.

XX

plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;

KW

animal feed product; thale cress; cell wall biosynthesis;

KW

nitrogen metabolism; carbon metabolism.

XX

Arabidopsis thaliana.

OS

WO2004035798-A2.

FN

29-APR-2004.

XX

20-OCT-2003; 2003WO-EP011658.

PR

18-OCT-2002; 2002EP-00079408.

XX

(CROP-) CROPDESIGN NV.

PA

Inze D, De Veylder L, Vlieghe K;

PI

WPI; 2004-348466/32.

XX

N-PSDB; ADN73862.

DR

XX


```
RESULT 8
ABB91288
ID ABB91288 standard; protein; 435 AA.
XX
XX
AC ABB91288;
XX
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 499.
XX
XX Herbicidal; plant; agriculture; herbicide.
XX
XX Arabidopsis thaliana.
XX
XX WO200210210-A2.
XX
XX PD 07-FEB-2002.
XX
XX PF 28-AUG-2001; 2001WO-EP009892.
XX
XX PR 28-AUG-2001; 2001WO-EP009892.
XX
XX PA (FARB ) BAYER AG.
XX
XX PI Tietjen K, Weidler M;
XX
XX DR WPI; 2002-269010/31.
XX
XX PT Identifying plant target proteins for herbicidally active compounds,
XX comprising aligning and comparing nucleic acid or amino acid sequences
XX from plant with nucleic acid or amino acid sequences from non-plant
XX organisms.
XX
XX PS Claim 5; SEQ ID NO 499; 261pp + Sequence Listing; English.
XX
XX The invention relates to identifying target proteins (ABB90790-ABB94016)
XX for herbicidally active compounds, comprising aligning and comparing
XX nucleic acid or amino acid sequences from plant with nucleic acid or
XX amino acid sequences from non-plant organisms using suitable search
XX parameters, where plant sequences having an E-value greater by a factor
XX of 3 than the E-value of most similar non-plant sequences are selected.
XX The polypeptides or nucleic acids encoding them are useful for
XX identifying modulators. The identified modulators are useful as
XX herbicides
XX
XX SQ Sequence 435 AA;
XX
XX Query Match 65.4%; Score 1438.5; DB 5; Length 435;
XX Best Local Similarity 67.4%; Pred. No. 1.3e-131;
XX Matches 267; Conservative 65; Mismatches 63; Indels 1; Gaps 1;
XX
XX QY 31 ASSLRFAVHRCRTARLEVKASGNTFGNYFQVATYGGSHGGVGVGCVISGCPPIPLTEAD 90
XX :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|:::
XX Db 30 SSPAVQISLRITQTRKNFQIQATGSSYGFHFRVSTFGSHGGVGGCIIDGCPPIPLTESD 89
XX
XX QY 91 LQVELDRRPGOSRITSTRKETDTCKILSGTHEGVTTCPIPLVIVPNTDQIGSDHREIAN 150
XX :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|:::
XX Db 90 LQFDLD-RPQGOSRITTRPRKETDTCKIRISGVSEGMTTCPIHVFVPNTDQGLDYSEMSV 148
XX
XX QY 151 VYRSHADATYDFKYGVRVAVGGGRSSGRKTVGRVAAGALPKKILKLCGLLEILSFVSKV 210
XX :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|:::
XX Db 149 AYRPSHADATYDMKYGVRSVQGGGRSSARETIGRVAPGLAKKILKQFAGTEILAYSQV 208
XX
XX QY 211 HOVLPEADYVGSVTLTQIESNIVRCPPDPEYAEKMDAIDRRVRGDSVGGVITCVARN 270
XX :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|:::
XX Db 209 HHVLPPEELVDHENLTLEQIENNVRCFNPPEYAEKMAAIDAVRTKGNVGGVWTCIVRN 268
XX
XX QY 271 VPRGIGSVFQKLESELAKMILSPASNGFEIGSGFAGTDLTGSHNDEFYMDKAGSVRT 330
XX :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|:::
XX Db 269 APRGLGTVPFQKLEAEAKACMSLPATKGFPGSGFAGTFTLTGLHNDEFFYDENGRI 328
XX
XX QY 331 RTNRSGGVQGGISNVEIVHFKVAFKPTPSIGVKQNTVSRERQNVELLARGRHDPCVAPRA 390
XX :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|:::
XX
329 RTNRSGGIQQGISNGEIIINRVAFKPTSTIGRKQNTVTRDKVETMIARGRHDPVVPRA 388
QY 391 VPVVESMAALVLMQDLMAHVAQCCEMPALNTALQEPV 426
XX :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|:::
XX Db 389 VPMVEAMVALVLDQLMAQYAQCHLFPINPELOEPL 424
XX
RESULT 9
AAY44895
ID AAY44895 standard; protein; 307 AA.
XX
XX AC AAY44895;
XX
XX DT 18-MAY-2000 (first entry)
XX
XX DE Rice chorismate synthase encoded by clone rls72.pk0029.g8.
XX
XX KW Rice cDNA clone rls72.pk0029.g8; chorismate synthase;
XX immunological screening; herbicide resistance; antibody; gene mapping;
XX rice.
XX
XX OS Oryza sativa.
XX
XX PN WO200005353-A2.
XX
XX PD 03-FEB-2000.
XX
XX PF 20-JUL-1999; 99WO-US016353.
XX
XX PR 21-JUL-1998; 98US-0093611P.
XX
XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX PI Cahoon RE, Falco SC, Pember SO;
XX
XX DR WPI; 2000-182687/16.
XX
XX N-PSDB; AAZ50440.
XX
XX PT New chorismate synthase polypeptides used to alter the level of the
XX enzyme and thus the level of aromatic to non aromatic amino acids in
XX transformed plants.
XX
XX PS Claim 6; Page 35-36; 39pp; English.
XX
XX The present sequence is rice chorismate synthase encoded by cDNA clone
XX designated rls72.pk0029.g8. Manipulating either the amount or activity of
XX this enzyme would affect the ratio of aromatic to non-aromatic amino
XX acids in plants. Chimeric gene comprising the cDNA operably linked to
XX regulatory sequences is used to transform host cells to alter the level
XX of expression of chorismate synthase. The gene and its products may be
XX used for immunological screening of cDNA expression libraries and to
XX create transgenic plants which may also be herbicide resistant. Synthetic
XX peptides derived from the gene are to raise antibodies, and used in
XX screening assays to identify inhibitors which may be useful as herbicides
XX
XX SQ Sequence 307 AA;
XX
XX Query Match 57.1%; Score 1257; DB 3; Length 307;
XX Best Local Similarity 81.9%; Pred. No. 4.6e-114;
XX Matches 245; Conservative 22; Mismatches 26; Indels 6; Gaps 2;
XX
XX QY 15 ARLAPRAIG----ALLEFAPASSSLRFAVHRCRTARLEVKASGNTFGNYFQVATYGGSHG 70
XX :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|:::
XX Db 11 ARVLPGGGGGGFRAPFESAPA--SLRFSVGRRAARLEVKASANVFNQVATYGGSHG 68
XX
XX QY 71 GGVGCVISGCPPIPLTEADLQVELDRRPGOSRITSTRKETDTCKILSGTHEGVTTCPIPL 130
XX :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|:::
XX Db 69 GGVGCVISGCPPIPLTEADQVELDRRPGOSRITTRPRKETDTCKILSGTHEGVTTCPIPL 128
XX
XX QY 131 ILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGVRVAVGGGRSSGRKTVGRVAAGAL 190
XX :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|:::
XX Db 129 IHVFVPNTDQGGDYSMAKAYRPSHADATYDFKYGVRVAVGGGRSSARETIGRVAAGAL 188
```


PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX
XX Claim 1; SEQ ID NO 2677; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 362 AA;

Query Match 54.2%; Score 1191.5; DB 8; Length 362;
Best Local Similarity 62.2%; Pred. No. 1.5e-107;
Matches 225; Conservative 57; Mismatches 77; Indels 3; Gaps 1;

QY 53 GNTFGNYFQVATYGESHGGVCGVISCPPRIPLTEADLOVELDRRPGQSRTSTKET 112
DB 2 GNTFGSLFRITTFGESHGGVGIIDGCPPLKLEISPEEIQVLDLRRRPGQSKITTPRKEA 61
QY 113 DTCKILSGTHEGVTGTPIILVNPNTDQIGSDHREIANVYRPSHADATYDFKYGVRVQ 172
DB 62 DQCEILSGVFEGKTGTPIILVNRKDSQDYNEMAVKYPSPHADATYEAQYGRWQ 121
QY 173 GRRSGRTYGRVAAAGALPKKILKLCGLEILSFVSKVHQVLPEDAVDYGSTVLEQIES 232
DB 122 GGRSARETIGRVAAGAIAKILAQFNGEIVAYVKSQDI---EATVDSNTVILEQVES 178
QY 233 NIVRCPDEYAEKMDAIDRVVRGDSVGVITCVARNVPLGSPVFDKLESLAKAML 292
DB 179 NIVRCPDECAKMTIERIDQVLKQDSTGGVVECAIRNAPKLGEPVFDKLEADLAKAM 238
QY 293 SIPASNGEISGFGAGTDLTGEHNDYFMDKAGSVRTNRSGVQGGISNVEIVFKV 352
DB 239 SLPATKGFSGGFGAGTLLTSQNDYLLDEAGEWRTNRSGVQGGISNGEPIINRI 298
QY 353 AFKPPSPGKQNTYRQRNVELLARHPDPCVAPRAVPVVEGMAALVLMQDLMAHVAQ 412
DB 299 AFKPTATIGEQKTVNSNIGETTLAAKGRHPDPCVLPRAVPVVEGMAALVLDLHLRFOAQ 358
QY 413 CE 414
DB 359 CK 360

RESULT 12
ADS29835
ID ADS29835 standard; protein; 362 AA.
XX
AC ADS29835;
XX

DT 02-DEC-2004 (first entry)

XX Bacterial polypeptide #18868.
DE
XX Recombinant DNA construct; transformed plant; improved plant property;
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX bacterial polypeptide.
OS Bacteria.
XX US2003233675-A1.
XX 18-DEC-2003.
XX 20-FEB-2003; 2003US-00369493.
XX 21-FEB-2002; 2002US-0360039P.
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 18868; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 362 AA;

Query Match 53.9%; Score 1186.5; DB 8; Length 362;
Best Local Similarity 60.6%; Pred. No. 4.8e-107;
Matches 220; Conservative 66; Mismatches 74; Indels 3; Gaps 1;

QY 53 GNTFGNYFQVATYGESHGGVCGVISCPPRIPLTEADLOVELDRRPGQSRTSTKET 112
DB 2 GNTFGSLFRITTFGESHGGVGIIDGCPPLKLEISPEEIQVLDLRRRPGQSKITTPRKEA 61
QY 113 DTCKILSGTHEGVTGTPIILVNPNTDQIGSDHREIANVYRPSHADATYDFKYGVRVQ 172

Db 62 DTCEILSGVYEGKTLGTPISILVRNKDTRPDQYDEMAQKRPESHADATYDAKYGRNQ 121
QY 173 GRRSGRKTGVRVAAGALPKKILKCGLEILSFVSKVHQVLPEDAVDYGSTLEQIES 232
Db 122 GGRSSARETIGRVAAGAIAKKILRQVANVEIGYVKRIKDL---EGVDPNTVTLDOVES 178
QY 233 NIVRCPDPEYAEKMDAIDRVRVRGDSVGVITCVARNVPRGLGSPVFDKLESELAKAML 292
Db 179 NIVRCPDGLADRMIELISQTGRQDSIGGVVCEVARNVPGKLGEPVFDKLEADIAPVM 238
QY 293 SPASNGFEIGSGFAGTDLTGSHNDNFYMDKAGSVRTNRSGGVQGGISNVEIVHFKV 352
Db 239 SLPASRGFEIGSGFAGTDLTGSHNDNFYMDKAGSVRTNRSGGVQGGISNVEIVHFKV 298
QY 353 AFKPTPSIGVKONTVSRERQNVELLARGHDPVAPRAVPVVSMAALVLMQMAHVAQ 412
Db 299 AFKPTATIRKEQKTVTREGEETLLAAKGRHDPVLPRAVPVMEAMVALVCDHLRHHGQ 358
QY 413 CEM 415
Db 359 CKV 361

RESULT 13

ADS41780

ID ADS41780 standard; protein; 362 AA.

XX ADS41780;

AC ADS41780;

DT 02-DEC-2004 (first entry)

DE Bacterial polypeptide #20210.

XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

OS Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO 'Y.

XX (HINK/) HINKLE G J.

XX (SLAT/) SLATER S C.

XX (CHEN/) CHEN X.

XX (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX Claim 1; SEQ ID NO 20210; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a

CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or by
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 362 AA;

Query Match 53.8%; Score 1182.5; DB 8; Length 362;

Best Local Similarity 60.6%; Pred. NO. 1.2e-106;

Matches 220; Conservative 65; Mismatches 75; Indels 3; Gaps 1;

QY 53 GNTFGNYFOVATYGESHGSGVGVITGCPRIPLTEADLOVELDRRRQSPITSTRKET 112

Db 2 GNIFGHLFRISTFGESHGSGVGVITGCPRIPLTEADLOVELDRRRQSPITSTRKET 61

QY 113 DTCKILSGTHEGVTGTPILVVPNTDQIGSDHREIANVYRPSHADATYDFKVGVRVQ 172

Db 62 DTCEILSGVYEGKTLGTPISILVRNKDTRPDQYDEMAQKRPESHADATYDAKYGRNQ 121

QY 173 GRRSGRKTGVRVAAGALPKKILKCGLEILSFVSKVHQVLPEDAVDYGSTLEQIES 232

Db 122 GGRSSARETIGRVAAGAIAKKILRQVANVEIGYVKRIKDL---EGVDPNTVTLDOVES 178

QY 233 NIVRCPDPEYAEKMDAIDRVRVRGDSVGVITCVARNVPRGLGSPVFDKLESELAKAML 292

Db 179 NIVRCPDPEYAEKMDAIDRVRVRGDSVGVITCVARNVPRGLGSPVFDKLESELAKAML 238

QY 293 SPASNGFEIGSGFAGTDLTGSHNDNFYMDKAGSVRTNRSGGVQGGISNVEIVHFKV 352

Db 239 SLPASRGFEIGSGFAGTDLTGSHNDNFYMDKAGSVRTNRSGGVQGGISNVEIVHFKV 298

QY 353 AFKPTPSIGVKONTVSRERQNVELLARGHDPVAPRAVPVVSMAALVLMQMAHVAQ 412

Db 299 AFKPTATIRKEQKTVTREGEETLLAAKGRHDPVLPRAVPVMEAMVALVCDHLRHHGQ 358

QY 413 CEM 415

Db 359 CKV 361

RESULT 14

AAG24019

ID AAG24019 standard; protein; 313 AA.

XX AAG24019;

XX AAG24019;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 27536.

XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX

PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0158293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 21-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161992P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.
 Query Match 51.7%; Score 1137; DB 3; Length 313;
 Best Local Similarity 70.9%; Pred. No. 2.7e-102;
 Matches 214; Conservative 42; Mismatches 46; Indels 0; Gaps 0;
 QY 125 VTGTPILVIVNTDQIGSHREIAVVRPSHADATYDFKYGVRVQGGRRSSGRKTVGR 184
 DB 1 MTGTPIHVFVNTDQIGSHREIAVVRPSHADATYDFKYGVRVQGGRRSSGRKTVGR 60
 QY 185 VAAGALPKKILKCGLEILSFVSKVHVLPEDAVDYGSVTLEQIESNIVRCPPDEYAE 244
 DB 61 VAFGALAKKILKQAGTEILAVSVQVHVLPEDAVDYGSVTLEQIESNIVRCPPDEYAE 120
 QY 245 KMIDAIDRVVRGDSVGGVITCVARNVRGLGSPVDFKLESELAKAMLSIPASNGFEIGS 304
 DB 121 KMIAIDAIVRTKSGNSVGGVITCVARNVRGLGTPVDFKLESELAKAMLSIPATKGFBEFS 180
 QY 305 FGAGTDLGSEHNDFFYMDKAGSVTRTRNRSVGGVGGISNVEIVHFKVAFKPTPSIGVKQ 364
 DB 181 FGAGTDLGSEHNDFFYMDKAGSVTRTRNRSVGGVGGISNVEIVHFKVAFKPTPSIGVKQ 240
 QY 365 NTVSRERQNVLLARGRHDPVAPRAVPVWESMAALVMDQLMAHVAQCEMFALNTALQE 424
 DB 241 NTVTRDKVETEMIAGRHDPVAPRAVPVWESMAALVMDQLMAHVAQCEMFALNTALQE 300
 QY 425 PV 426
 DB 301 PL 302

RESULT 15
 ID AAY44892
 XX AAY44892 standard; protein; 257 AA.
 AC AAY44892;

XX 18-MAY-2000 (first entry)
 XX Wheat chorismate synthase encoded by clone wreln.pk0094.e6.
 KW Wheat cDNA clone wreln.pk0094.e6; chorismate synthase;
 KW immunological screening; herbicide resistance; antibody; gene mapping;
 wheat.
 XX Triticum aestivum.
 XX WO200005353-A2.
 XX 03-FEB-2000.
 XX 20-JUL-1999; 99WO-US016353.
 XX 21-JUL-1998; 98US-0093611P.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX Cahoon RE, Falco SC, Pember SO;
 XX WPI; 2000-182687/16.
 XX N-PSDB; AAZ50437.
 XX New chorismate synthase polypeptides used to alter the level of the
 enzyme and thus the level of aromatic to non aromatic amino acids in
 transformed plants.
 XX Claim 6; Page 31-32; 39pp; English.
 XX The present sequence is wheat chorismate synthase encoded by cDNA clone
 designated wreln.pk0094.e6. Manipulating either the amount or activity of
 this enzyme would affect the ratio of aromatic to non-aromatic amino
 acids in plants. Chimeric gene comprising the cDNA operably linked to
 regulatory sequences is used to transform host cells to alter the level
 of expression of chorismate synthase. The gene and its products may be
 used for immunological screening of cDNA expression libraries and to
 create transgenic plants which may also be herbicide resistant. Synthetic
 peptides derived from the gene are to raise antibodies, and used in
 screening assays to identify inhibitors which may be useful as herbicides
 SQ Sequence 257 AA;
 Query Match 48.5%; Score 1066; DB 3; Length 257;
 Best Local Similarity 82.7%; Pred. No. 1.8e-95;
 Matches 201; Conservative 27; Mismatches 15; Indels 0; Gaps 0;
 QY 186 AAGALPKKILKCGLEILSFVSKVHVLPEDAVDYGSVTLEQIESNIVRCPPDEYAEK 245
 DB 3 AAGAVAKKILKCGVEILAFVSKVHVLPEDAVDYGSVTLEQIESNIVRCPPDEYAEK 62
 QY 246 MIDAIDRVVRGDSVGGVITCVARNVRGLGSPVDFKLESELAKAMLSIPASNGFEIGSG 305
 DB 63 MIDAIDRVVRGDSVGGVITCVARNVRGLGSPVDFKLESELAKAMLSIPASNGFEIGSG 122
 QY 306 FAGTDLGSEHNDFFYMDKAGSVTRTRNRSVGGVGGISNVEIVHFKVAFKPTPSIGVKQ 365
 DB 123 FAGTDLGSEHNDFFYMDKAGSVTRTRNRSVGGVGGISNVEIVHFKVAFKPTPSIGVKQ 182
 QY 366 TVSRERQNVLLARGRHDPVAPRAVPVWESMAALVMDQLMAHVAQCEMFALNTALQEP 425
 DB 183 TVTRDHEDIELTRGRHDPVAPRAVPVWESMAALVMDQLMAHVAQCEMFALNTALQEP 242
 QY 426 VGS 428
 DB 243 IGS 245

Search completed: August 25, 2005, 19:30:06
 Job time : 80 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 19:25:10 ; Search time 19 seconds
(without alignments)
2172.473 Million cell updates/sec

Title: US-10-677-179-8
Perfect score: 2200
Sequence: 1 MTTVPKQQVAHSRARLAPR.....VAQCEMFALNTALQBPVGSF 429

Scoring table: BLOSUM62
dapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1541	70.0	431	2 S40409	chorismate synthas
2	1527	69.4	440	2 S40410	chorismate synthas
3	1486.5	67.6	447	1 A41197	chorismate synthas
4	1438.5	65.4	435	2 C96526	probable chorismat
5	1191.5	54.2	362	2 A49316	chorismate synthas
6	1186.5	53.9	362	2 AC1906	chorismate synthas
7	951	43.2	395	2 T41268	chorismate synthas
8	938	42.6	376	1 S17246	chorismate synthas
9	930	42.3	432	2 T46725	chorismate synthas
10	924	42.0	432	2 T51020	chorismate synthas
11	862	39.2	361	2 AG0335	chorismate synthas
12	860	39.1	361	1 SYECKR	chorismate synthas
13	856	38.9	361	2 E91030	chorismate synthas
14	856	38.9	361	2 F85874	chorismate synthas
15	850	38.6	372	1 A55510	chorismate synthas
16	841	38.2	354	2 H84940	chorismate synthas
17	840.5	38.2	357	1 G64053	chorismate synthas
18	840	38.2	361	2 F82115	chorismate synthas
19	833	37.9	361	1 SYE8KR	chorismate synthas
20	830	37.7	361	2 AF0804	chorismate synthas
21	817	37.1	366	2 A81822	chorismate synthas
22	806	36.6	366	2 B81055	chorismate synthas
23	798.5	36.3	378	2 F64446	chorismate synthas
24	787.5	35.8	360	2 F69333	chorismate synthas
25	779.5	35.4	363	2 B83436	chorismate synthas
26	771	35.0	364	2 AD3440	chorismate synthas
27	766.5	34.8	374	2 A69200	chorismate synthas
28	764	34.7	365	2 AD2669	chorismate synthas
29	764	34.7	368	2 B97451	chorismate synthas

ALIGNMENTS

RESULT 1

S40409

Chorismate synthase (EC 4.2.3.5) 2 precursor - tomato
C:Species: Lycopersicon esculentum (tomato)
C:Date: 19-May-1994 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S40409; S38471
R:Goerlach, J.; Schmid, J.; Amrhein, N.
Plant Mol. Biol. 23, 707-716, 1993
A:Title: Differential expression of tomato (Lycopersicon esculentum L.) genes encoding S40409
A:Reference number: S40409; MUID:94072719; PMID:8251624
A:Accession: S40409
A:Molecule type: mRNA
A:Residues: 1-431 <GO>
A:Cross-references: UNIPROT:Q42885; EMBL:Z21791; NID:G410483; PIDN:CAA79854.1; PID:G4104

Query Match	70.0%;	Score	1541;	DB 2;	Length	431;			
Best Local Similarity	72.7%;	Pred. No.	2.9e-108;						
Matches	287;	Conservative	55;	Mismatches	53;	Indels	0;	Gaps	0;
QY	32	SSSLRFVAVHRCRTARLEVKASGNTFGNTFQVATYGESHGGGVCVTSGCCPPRIPLTEADL	91						
Db	29	SSNLRFPTHRSPKLEIQAAAGNTFGNTFQVATYGESHGGGVCVTSGCCPPRIPLTEADL	88						
QY	92	QVELDRRRPGQSRIITSTKRTDTCILSGTHGVTGTPILVIVPNTDQIGSDHREIANV	151						
Db	89	QVELDRRRPGQSRIITSTKRTDTCILSGTADGLATGSPIKVEVPNTDQRGNDYSMSLA	148						
QY	152	YRPSHADATYDFKYGVRAVQGGSSGRKTVGRVAAGALPKKILKLCGLLEILSFVSKVH	211						
Db	149	YRPSHADATYDFKYGVRAVQGGSSGRKTVGRVAAGALPKKILKLCGLLEILSFVSKVH	208						
QY	212	QVLPEDAVDYGSVTLEQIESNIVRCPDPEYAEKMDAIDRVVRGDSVGGVITCVARNV	271						
Db	209	NVLPEDLVNDQIVTLEQIESNIVRCPNPEYAEKMDAIDRVVRGDSVGGVITCVARNV	268						
QY	272	PRGLGSPVDFKLESLAKAMLSIPASNGFEIGSGFAGTDLTGSEHNDFEYMDKAGSVRTR	331						
Db	269	PRGLGTPVDFKLESLAKACMSLPATKGFEGSGFAGTGMTSEHNDFEYMDKAGSVRTR	328						
QY	332	TNRSGGVGGISNVEIVHFKVAPKTPSIGVKQNTVSRQNVLLARGRHDPCCVAPRAV	391						
Db	329	TNRSGGIQGGISNGEILNWRVAFKSTIARQHTVSRDKHETELIARGRHDPCCVAPRAV	388						
QY	392	PVVESSMAALVIMDQLMAHVAQCEMFALNTALQEPV	426						
Db	389	PMVEAMVALVLDQIMTQVACQMLFPVNLTLQEPL	423						

chorismate synthas
chorismate synthas
chorismate synthas
chorismate synthas
chorismate synthas
probable chorismat
chorismate synthas
chorismate synthas
chorismate synthas
chorismate synthas
chorismate synthas
chorismate synthas

RESULT 2

S40410
Chorismate synthase (BC 4.2.3.5) 1 precursor - tomato
C:Species: Lycopersicon esculentum (tomato)
C>Date: 19-May-1994 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S40410; S38470
R:Goerlach, J.; Schmid, J.; Amrhein, N.
Plant Mol. Biol. 23, 707-716, 1993
A:Title: Differential expression of tomato (Lycopersicon esculentum L.) genes encoding s
A:Reference number: S40409; MUID:94072719; PMID:8251624
A:Accession: S40410
A:Molecule type: mRNA
A:Residues: 1-440 <GO>
A:Cross-references: UNIPROT:Q42884; EMBL:Z21796; NID:G410481; PIDN:CAA79859.1; PID:G4104
C:Genetics:
A:Genome: nuclear
C:Superfamily: chorismate synthase
C:Keywords: aromatic amino acid biosynthesis; carbon-oxygen lyase; chloroplast; phosphon
F:1-54/Domain: transit peptide (chloroplast) #status predicted <TNP>
F:55-440/Product: chorismate synthase 1 #status predicted <MAT>

Query Match 69.4%; Score 1527; DB 2; Length 440;
Best Local Similarity 67.4%; Pred. NO. 3.4e-107;
Matches 288; Conservative 64; Mismatches 73; Indels 2; Gaps 1;
Qy 4 VPKQQAHAHRAR--LAPRATGALLEFAPASSLRFAVHRCRTARLEVKASGNTFGNYFQ 61
Db 5 VPTKQFVGASSSSDITGSSRLVSLQLPSKFSNPHLPSPQSLKRLTQAAGSTFGNYFR 64
Qy 62 VATGESGGGVCVIGSCPPRIPLTEADLOVELDRRPPGOSRITSTRKETDTCKILSGT 121
Db 65 VTTGESGGGVCCTIDCCPLPLPSESDMQVELDRRPPGOSRITSTRKETDTCKILSGT 124
Qy 122 HEGVTTGFPILVTPNTQIGSDHREIANVYRPSHADATYFKYGVRAVQGGSSGRKT 181
Db 125 ADGLTTGSPKVEPNTDQGNDSYSEMSLAYRPSHADATYFKYGVRSVQGGSSARET 184
Qy 182 VGRVAAGALPKILKLCGLBILFVSKVHQVLPEDAVYGVSTVLEQIESNIVRCDDPE 241
Db 185 IGRVAAGAVAKILKLYGAEVLAVYSQVHQVLPEDLIDHQNTLEQIESNIVRCDDPE 244
Qy 242 YAEKMDAIDRVVRGDSVGVITCVARNVRGLGSPVFDKLESELAKAMLSIASNGFE 301
Db 245 YAEKMDAIDAVVRGDSVGVVITCVARNVRGLGTPVFDKLESELAKAMLSLPATKGE 304
Qy 302 IGSFAGTDLTGSEHNDFYMDKAGSVRTRNRSVGQGISNVEIVHFVKFAPKTPPSIG 361
Db 305 FGSFAGTFTMTGSEHNDFYMDKAGSVRTRNRSVGQGISNVEIVHFVKFAPKTPPSIG 364
Qy 362 VKQNTVSRERQNVELLARHDPVAPRAVPVVSMAALVMDQLMAHVAQCEMPALNTA 421
Db 365 RKQQTVRDRKHETELIARHDPVAPRAVPVVSMAALVMDQLMAHVAQCEMPALNTA 424
Qy 422 LQEPVGS 428
Db 425 LQEPVGS 431

RESULT 3

A41197
Chorismate synthase (BC 4.2.3.5) precursor [validated] - pink corydalis
C:Species: Corydalis sempervirens (pink corydalis)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: A41197; S13267; S15715
R:Schaller, A.; Schmid, J.; Leiberger, U.; Amrhein, N.
J. Biol. Chem. 266, 21434-21438, 1991
A:Title: Molecular cloning and analysis of a cDNA coding for chorismate synthase from th
A:Reference number: A41197; MUID:92042037; PMID:1718979
A:Accession: A41197
A:Status: preliminary
A:Molecule type: mRNA

A:Residues: 1-447 <SCH>

A:Cross-references: UNIPROT:P27793; EMBL:X60544; NID:gl8255; PIDN:CAA43034.1; PID:gl8255
R:Schaller, A.; Windhofer, V.; Amrhein, N.
Arch. Biochem. Biophys. 282, 437-442, 1990
A:Title: Purification of chorismate synthase from a cell culture of the higher plant Cory
A:Reference number: S13267; MUID:91053166; PMID:2146922
A:Accession: S13267

A:Molecule type: protein

A:Residues: "X", 307-320 <SCW>

C:Genetics:

A:Genome: nuclear

C:Complex: homodimer

C:Function:

A:Description: EC 4.6.1.4 [validated, MUID:91053166]; catalyzes the formation of chorism
C:Superfamily: chorismate synthase
C:Keywords: aromatic amino acid biosynthesis; carbon-oxygen lyase; chloroplast; phosphon
F:1-57/Domain: transit peptide (chloroplast) #status predicted <TNP>
F:58-447/Product: chorismate synthase #status predicted <MAT>

Query Match 67.6%; Score 1486.5; DB 1; Length 447;
Best Local Similarity 70.4%; Pred. NO. 3.8e-104;
Matches 280; Conservative 60; Mismatches 55; Indels 3; Gaps 2;
Qy 31 ASSSLRPAVHRCRTA--RLEVKASGNTFGNYFOVATYGESHGSGVGCVIGSCPPRIPLTE 88
Db 37 SNQSVQISVPR-QTAPLKLVVQASGSGFKVQVTTYGESHGSGVGCVIDGCPRPPISE 95
Qy 89 ADIQVELDRRRPGOSRITSTRKETDTCKILSGTHEGVTGTPILVIVPNTDQIGSDHREI 148
Db 96 ADIQVELDRRRPGOSRITSTRKETDTCKILSGVADGFTTGTGSIHISVPNTDQIGSDHREI 155
Qy 149 ANVYRSHADATYDFKYGVRAVQGGSSGRKTVGRVAAGALPKILKLCGLBILSFVS 208
Db 156 AKYRPSHADATYDFKYGVRSVQGGSSARETIGRVAAGALAKKILKAYAGTEVLAYS 215
Qy 209 KVHQVLPEDAVYGVSTVLEQIESNIVRCDDPEYAEKMDAIDRVVRGDSVGVITCV 268
Db 216 QAHKVLPEGLVDHETLSLEQIESNIVRCDDPEYAEKMDAIDRVVRGDSVGVITCV 275
Qy 269 RNVPRLGSPVFDKLESELAKAMLSIPASNGFEIGSGFAGTDLTGSHNDFYMDKAGSV 328
Db 276 RNVPRLGSPVFDKLESELAKAMLSIPATKFGFGSGFTTGTGSEHNDFYTDENGRI 335
Qy 329 RTNRSVGQGISNVEIVHFVKFAPKTPPSIGVKQNTVSRERQNVELLARHDPVAP 388
Db 336 RTNRSVGQGISNVEIVHFVKFAPKTPPSIGVKQNTVSRERQNVELLARHDPVAP 395
Qy 389 RAVPVVSMAALVMDQLMAHVAQCEMPALNTALQEPV 426
Db 396 RAVPVVSMAALVMDQLMAHVAQCEMPALNTALQEPV 433

RESULT 4

C96526

probable chorismate synthase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C:Accession: C96526

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, R.

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: C96526

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-435 <STO>

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-362 <KAN>
A:Cross-references: EMBL:D90906; GB:AB001339; NID:g1652492; PIDN:BAAL7415.1; PID:g1652492
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: aroC
C:Superfamily: chorismate synthase
C:Keywords: carbon-oxygen lyase; phosphorus-oxygen lyase

	Query Match	54.2%;	Score 1191.5;	DB 2;	Length 362;
	Best Local Similarity	62.2%;	Pred. No. 4.5e-82;		
	Matches 225;	Conservative 57;	Mismatches 77;	Indels 3;	Gaps 1;
QY	53	GNTFGNYFOVATYGSHGGGVCVIGCGPPRIPLTEADLQVELDRRRPGOSRITSTRKET	112		
DB	2	GNTFGSLFRITTFGSHGGGCVIIDGCPRLISPEEIQVLDLDRRPQGSKITTPRKEA	61		
QY	113	PTCKILSGTHEGVTGTTPILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGVRVQG	172		
DB	62	DQCEILSGVFEGKTLGTPIAILVRNKDARSQDYNEMAVKYPESHADATYEAKYGIRNWQG	121		
QY	173	GGSSGRKTVGVAAGALPKKILKCGLEILSPVSKHQVVLPEDAVDVGSVTLEQIES	232		
DB	132	GGSSARETIGVAAGAIKALQAPNGVEIVAVKSIQDI---EATVDSNTVTLEQVES	178		
QY	233	NIVRCPDPEYASKMIDAIIDRVVRGDSVGVITCVARNVPRGLGSPVFDKLESLAKAML	292		
DB	179	NIVRCPDECAKMLERIDQVLRQKDSIGGVVECATRNAPKGLGEPVFDKLEADLAKAMM	238		
QY	293	SIPASNGFEIGSGFAGTDLTGSEHNDEFYMDKAGSVTRTRNSGGVQGGISNVIVHFKV	352		
DB	239	SLPATKGFEGSGFAGTLLTGSHQNDYYLDEAGWRTRTRNSGGVQGGISNGEPIIMRI	298		
QY	353	AFKPTPSTGVKQNTVSRERQNVELLARGHDPVAPRAVPVSVESMAALVLMQDLMAHVAQ	412		
DB	299	AFKPTATIGQEOKTVSNIGETTLAAGRHDPVLPRAVPMVEAMALVLCDDLHLLRFQAAQ	358		
QY	413	CE 414			
DB	359	CK 360			

RESULT 6
AC1906
chorismate synthase [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AC1906
R,Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shampo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AC1906
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-362 <KUR>
A:Cross-references: UNIPROT:Q8YTP9; GB:BA000019; PIDN:BA72754.1; PID:g17130142; GSPDB:G
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all0797
C:Superfamily: chorismate synthase

	Query Match	53.9%;	Score 1186.5;	DB 2;	Length 362;
	Best Local Similarity	60.6%;	Pred. No. 1.1e-81;		
	Matches 220;	Conservative 66;	Mismatches 74;	Indels 3;	Gaps 1;
QY	53	GNTFGNYFOVATYGSHGGGVCVIGCGPPRIPLTEADLQVELDRRRPGOSRITSTRKET	112		
DB	2	GNTGHLFRITTFGSHGGGCVIIDGCPPLISPEEIQLELDLDRRPQGSKITTPRKEA	61		

QY 406 LMAHVAQ 412
Db 362 LLIQXAR 368

RESULT 9

T46725
chorismate synthase (EC 4.2.3.5) / flavin reductase, NADPH-dependent [validated] - Neuro
N:Contains: chorismate synthase (EC 4.6.1.4); flavin reductase, NADPH-dependent
C:Species: Neurospora crassa
C>Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004
C:Accession: T46725
R:Henstrand, J.M.; Amrhein, N.; Schmid, J.
J. Biol. Chem. 270, 20447-20452, 1995
A:Title: Cloning and Characterization of a Heterologously Expressed Bifunctional Chorism
A:Reference number: 208450; MUID:95386486; PMID:7657620
A:Accession: T46725
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-434 <HEN>
A:Cross-references: UNIPROT:Q12640; EMBL:U25818; NID:g976374; PIDN:AAC49056.1; PID:g9763

C:Function:

A:Description: EC 4.6.1.4 [validated, MUID:95386486]
A:Note: N. crassa chorismate synthase (CS) is a bifunctional enzyme, possessing also an
A:Note: CS and FT activity are encoded by a single domain
C:Function: <PRE>
A:Description: Flavin reductase [validated, MUID:95386486]
A:Note: N. crassa chorismate synthase (CS) is a bifunctional enzyme, possessing also an
A:Note: CS and FT activity are encoded by a single domain
C:Superfamily: chorismate synthase
C:Keywords: aromatic amino acid biosynthesis; carbon-oxygen lyase; phosphorus-oxygen ly

Query Match 42.3%; Score 930; DB 2; Length 432;

Best Local Similarity 48.9%; Pred. No. 2.7e-62;
Matches 201; Conservative 64; Mismatches 110; Indels 36; Gaps 7;

QY 54 NTFGNYFOVATYGESHGGVGVISGCPRIPLTEADLOVELDRRPPGOSRITSTRKETD 113
Db 2 STFGHYFRVTYTGESHCKSVGCVIVDGVPGMELTDDIQPMTRRRPGOSAITTFRDEKD 61
QY 114 TKILSGTHEGVTGTPILVIVPNTDQIGSDH-REIANVY-RPSHADATYDFKYGVRVQ 171
Db 62 RVIIQSGTEFGVTGTPIGLVMNEDQPKDYGNKTMIDYPRSHADWTYLEKYGVKASS 121
QY 172 GGRSSGRKTGVRVAAGALPKILKLCGLLEILSFVSKVHQVVL-----PED 218
Db 122 GGRSSARETIGRVAAGAIAEKYLKPRYGVVEIVAFVSSVSGSEHLFPPTAEHPSPSTNPEF 181

QY 219 AVDYGSVTLLEQIESNI-VRCPEPEAEKMDAIDRVVRGDSVGVVITCVARNVPRGLGS 277
Db 182 LKLVNSITRETVDSPFLVPCPDPAENKRMEDLITFRDNHDSIGGVITCVIRNVPGLGE 241
QY 278 PVFDKLESELAKAMLSIPASNGFEIGSGFAGTDLTGSEHNDEFY-----MDKAGS 327
Db 242 PAFDKLEAMLAHMLSIPTATKFEVSGGCGEVPGSIHNDPFSVAENTEIPPSVAASGA 301
QY 328 VR-----TRTNRSGVQGGISNVEIVHFVKAFKPTPSIGVKONTVSRERQNVLL- 377
Db 302 ARNGIPRPLTKTNTFSGGIQGGISNGAPIYFRVGFKPAATIGQETATYDGTSGVLA 361

QY 378 ARGHRDPCVAPRAVPVVESSMAALVMDQLMAHVAOCEMFALNTALQEPVGS 428
Db 362 AKGRHDPSPVPRAPVIVEAMALVMDAVLAHEARVTAKSLPPLKQTINS 412

RESULT 10

T51020
chorismate synthase/flavin reductase, NADPH-dependent [imported] - Neurospora crassa
N:Alternate names: protein B7F21.10
C:Species: Neurospora crassa
C>Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 02-Sep-2000
C:Accession: T51020

R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, July 2000
A:Reference number: Z25286
A:Accession: T51020
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-432 <SCH>
A:Cross-references: EMBL:AL389901; GSPDB:GN00116; NCSP:B7F21.10
A:Experimental source: BAC clone B7F21; strain OR74A
C:Genetics:
A:Gene: NCSP:B7F21.10
A:Map position: 6
A:Introns: 12/2; 92/1
C:Superfamily: chorismate synthase

Query Match 42.0%; Score 924; DB 2; Length 432;

Best Local Similarity 48.9%; Pred. No. 7.6e-62;
Matches 201; Conservative 63; Mismatches 111; Indels 36; Gaps 7;

QY 54 NTFGNYFOVATYGESHGGVGVISGCPRIPLTEADLOVELDRRPPGOSRITSTRKETD 113
Db 2 STFGHYFRVTYTGESHCKSVGCVIVDGVPGMELTDDIQPMTRRRPGOSAITTFRDEKD 61
QY 114 TKILSGTHEGVTGTPILVIVPNTDQIGSDH-REIANVY-RPSHADATYDFKYGVRVQ 171
Db 62 RVIIQSGTEFGVTGTPIGLVMNEDQPKDYGNKTMIDYPRSHADWTYLEKYGVKASS 121
QY 172 GGRSSGRKTGVRVAAGALPKILKLCGLLEILSFVSKVHQVVL-----PED 218
Db 122 GGRSSARETIGRVAAGAIAEKYLKLAYGVVEIVAFVSSVSGSEHLFPPTAEHPSPSTNPEF 181

QY 219 AVDYGSVTLLEQIESNI-VRCPEPEAEKMDAIDRVVRGDSVGVVITCVARNVPRGLGS 277
Db 182 LKLVNSITRETVDSPFLVPCPDPAENKRMEDLITFRDNHDSIGGVITCVIRNVPGLGE 241
QY 278 PVFDKLESELAKAMLSIPASNGFEIGSGFAGTDLTGSEHNDEFY-----MDKAGS 327
Db 242 PAFDKLEAMLAHMLSIPTATKFEVSGGCGEVPGSIHNDPFSVAENTEIPPSVAASGA 301
QY 328 VR-----TRTNRSGVQGGISNVEIVHFVKAFKPTPSIGVKONTVSRERQNVLL- 377
Db 302 ARNGIPRPLTKTNTFSGGIQGGISNGAPIYFRVGFKPAATIGQETATYDGTSGVLA 361

QY 378 ARGHRDPCVAPRAVPVVESSMAALVMDQLMAHVAOCEMFALNTALQEPVGS 428
Db 362 AKGRHDPSPVPRAPVIVEAMALVMDAVLAHQAHAKSLPPLKQTINS 412

RESULT 11

AG0335
chorismate synthase (EC 4.2.3.5) [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AG0335
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
i, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AG0335

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-361 <KUR>
A:Cross-references: UNIPROT:Q82D41; GB:AL590842; PIDN:CAC92990.1; PID:g15980729; GSPDB:8

Query Match

Best Local Similarity 48.3%; Pred. No. 2.8e-57;
Matches 175; Conservative 65; Mismatches 110; Indels 12; Gaps 3;

QY 52 SGNTRGNFYOVATYGESHGCGVCGVISCPCPRIPILTEADLQVELDRRPPGQSRTSTRKE 111
DB 2 AGNSIGQFRVTTTFGESHGIALGCIIDGVPPGIPITEADIQDLDRRPPGTSRYTTQRR 61

QY 112 TDTCKILSGTHEGVTTGTPILVIVPNTDQIGSDHREIANVVRPSHADATYDFKYGVRVQ 171
DB 62 LDQVRLSGVFGVTTGTSIGMIENTDORSQDYSIAIKDVFPRGHADTYTEQYGLRDYR 121

QY 172 GGRSSGRKTVGRVAAGALPKKILKLCGLEILSFVSKVHQVLPEDAVDYGSVTLQIE 231
DB 122 GGRSSSARETAMRVAAGAIAKKYLAQKFGVQVGRYLAQMGDVC--DLLDW----DLVE 174

QY 232 SNIVRCPPEYAEKMDAIDRVRVSGDSVGGVITCVARNVPRGIGSPVFDKLESLAKAM 291
DB 175 QNPFFCDDPD---KIDALDELMLALKKEDSIGAKTVTVASGVPAAGLGEVFDRLDADI 234

QY 292 LSIPASNGPEIGSGFAGTDLTGSEHNDEFYMDKAGSVTRTRNRSQVGGISNVEIVHFK 351
DB 235 MSINAVKVEIGDGFVAVTVKRGSENRDEITPQ-----GFQSNHAGIILGGISSGQPVVAH 289

QY 352 VAFKTPPSIGVKQNTVSRERQNVLLARGHDPVAPRAVPVSVESMAALVMDQLMAHVA 411
DB 290 TALKEPTSSIMVPGQITNRQGEAVEMVTRGRHDPVCGIRAVPIAEAMMALVMDHLRQRA 349

QY 412 QC 413
DB 350 QC 351

RESULT 12
SYECKER
chorismate synthase (EC 4.2.3.5) - Escherichia coli (strain K-12)
N:Alternate names: 5-enolpyruvylshikimate-3-phosphate phospholiyase
C:Species: Escherichia coli
C:Date: 31-Dec-1990 #sequence revision 05-Dec-1997 #text_change 09-Jul-2004
R:Blattner, P.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: G65005
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-361 <BLAT>
A:Cross-references: UNIPROT:P12008; GB:AE000321; GB:U00096; NID:g1788659; PIDN:AACT53389.
A:Experimental source: strain K-12, substrain MG1655
R:White, P.J.; Millar, G.; Coggins, J.R.
Biochem. J. 251, 313-322, 1988
A:Title: The overexpression, purification and complete amino acid sequence of chorismate
A:Reference number: S00509; MUID:88293429; PMID:2969724
A:Accession: S00509
A:Molecule type: DNA
A:Residues: 1-334; NAGDRFNGSPVTATGAKCRCD' <WHI>
A:Cross-references: EMBL:M33021
A:Accession: S28890
A:Molecule type: protein
A:Residues: 1-31 <WHI2>
R:Charles, I.G.; Lamb, H.K.; Pickard, D.; Dougan, G.; Hawkins, A.R.
J. Gen. Microbiol. 136, 353-358, 1990
A:Title: Isolation, characterization and nucleotide sequences of the *aroC* genes encoding
A:Reference number: JY0091; MUID:90218018; PMID:2182772
A:Accession: JY0091
A:Molecule type: DNA
A:Residues: 1-361 <CHA>
A:Cross-references: GB:M27714; NID:g145357; PIDN:AAA23487.1; PID:g145358
A:Experimental source: isolate C600
A:Note: the authors translated the codon TCG for residue 171 as Arg, CAA for residue 175
3 as Pro, AAA for residue 206 as Ile, GAC for residue 229 as Thr, GAC for residue 266 as
C:Genetics:
A:Gene: *aroC*

A:Map position: 51 min
C:Superfamily: chorismate synthase
C:Keywords: aromatic amino acid biosynthesis; carbon-oxygen lyase; FMN; homotetramer; ph
P:2-356/Product: chorismate synthase #status experimental <MAT>

Query Match 39.1%; Score 860; DB 1; Length 361;
Best Local Similarity 48.8%; Pred. No. 3.9e-57;
Matches 178; Conservative 61; Mismatches 106; Indels 20; Gaps 5;

QY 52 SGNTRGNFYOVATYGESHGCGVCGVISCPCPRIPILTEADLQVELDRRPPGQSRTSTRKE 111
DB 2 AGNTIGQLFRVTTTFGESHGIALGCIIDGVPPGIPILTEADLQHLDRRPPGTSRYTTQRR 61

QY 112 TDTCKILSGTHEGVTTGTPILVIVPNTDQIGSDHREIANVVRPSHADATYDFKYGVRVQ 171
DB 62 PDQVKILSGVFGVTTGTSIGLIENTDORSQDYSIAIKDVFPRGHADTYTEQYGLRDYR 121

QY 172 GGRSSGRKTVGRVAAGALPKKILKLCGLEILSFVSKVHQVLPEDAVDYGSVTLQIE 231
DB 122 GGRSSSARETAMRVAAGAIAKKYLAERFGEIRGCLTQMGI--PLDIKDW-----SQVE 174

QY 232 SNIVRCPPEYAEKMDAIDR---VRVSGDSVGGVITCVARNVPRGIGSPVFDKLESEL 287
DB 175 QNPFFCDDPD---KIDALDELMLALKKEDSIGAKTVTVASGVPAAGLGEVFDRLDADI 230

QY 288 AKAMLSIPASNGFEIGSGFAGTDLTGSEHNDEFYMDKAGSVTRTRNRSQVGGISNVEI 347
DB 231 AHALMSINAVKVEIGDGFVAVTVKRGSENRDEITPQ-----GFQSNHAGIILGGISSGQ 285

QY 348 VHFVAFKTPPSIGVKQNTVSRERQNVLLARGHDPVAPRAVPVSVESMAALVMDQLM 407
DB 286 IIAHMAUKPTSSITVPGRTINRFEGEEMITKGRHDPVCGIRAVPIAEAMMALVMDHL 345

QY 408 AHVAQ 412
DB 346 QRQAQ 350

RESULT 13
E91030
chorismate synthase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 050995;
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: E91030
R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: E91030
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-361 <HAY>
A:Cross-references: UNIPROT:Q8XC04; GB:BA000007; PIDN:BA836636.1; PID:g13362683; GSPDB:G
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: EC83213
C:Superfamily: chorismate synthase

Query Match 38.9%; Score 856; DB 2; Length 361;
Best Local Similarity 48.5%; Pred. No. 7.8e-57;
Matches 177; Conservative 62; Mismatches 106; Indels 20; Gaps 5;

QY 52 SGNTRGNFYOVATYGESHGCGVCGVISCPCPRIPILTEADLQVELDRRPPGQSRTSTRKE 111
DB 2 AGNTIGQLFRVTTTFGESHGIALGCIIDGVPPGIPILTEADLQHLDRRPPGTSRYTTQRR 61

QY 112 TDTCKILSGTHEGVTTGTPILVIVPNTDQIGSDHREIANVVRPSHADATYDFKYGVRVQ 171
DB 62 PDQVKILSGVFGVTTGTSIGLIENTDORSQDYSIAIKDVFPRGHADTYTEQYGLRDYR 121

QY 172 GGRSSGRKTVGRVAAGALPKKILKLCGLEILSFVSKVHQVLPEDAVDYGSVTLQIE 231
DB 122 GGRSSSARETAMRVAAGALPKKILKLCGLEILSFVSKVHQVLPEDAVDYGSVTLQIE 231

Db 122 GGRSSARETAMRVAAGAAIAKYLAEKFGIEINGCLTQMGGDI--PLEIKDW-----SQVE 174

QY 232 SNIVRCPPDEYAEKMDAIDR----VRVRGDSVGGVITTCVARNVPRGLSGSPVFDKLESEL 287

Db 175 QNPFPCPPD-----KIDALDELMRALKKEGDSIGAKVTIVVAGVPAGLGEVFDRLDADI 230

QY 288 AKAMLSIPASNGFEIGSGFAGTDLTGSEHNDEFYMDKAGSVRTRTNRSQVGGVGSINVEI 347

Db 231 AHALMSINAVKVGVEIGDGDVVALRGSQNRDEITKD-----GFQSNHAGGILGGISSGQQ 285

QY 348 VHFVKAFKTPPSIGVKQNTVSRERQNVELLARGRHDPVAPRAVPVSVESMAALVMDQLM 407

Db 286 IIAHMAKPTSSITVPGRTINRFGEEVEMITKGRHDPVCGIRAVPIAEAMLAIVLMDHLL 345

QY 408 AHVAQ 412

Db 346 RQRAQ 350

RESULT 14

F85874

Chorismate synthase [imported] - Escherichia coli (strain O157:H7, substrain EDL933)

C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C:Accession: F85874

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouis, K.; Apodaca,
Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: F85874

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-361 <STO>

A:Cross-references: UNIPROT:Q8XCQ4; GB:AE005174; NID:g12516691; PIDN:AA657458.1; GSPDB:G

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: aroC

C:Superfamily: chorismate synthase

Query Match 38.9%; Score 856; DB 2; Length 361;
Best Local Similarity 48.5%; Pred. No. 7.8e-57;
Matches 177; Conservative 62; Mismatches 106; Indels 20; Gaps 5;

QY 52 SGNRTGNFQVATYGESHGCGVGVCGPPIPLTEADLQVELDRRRPGQSRTITRKE 111

Db 2 AGNTIGQLFRVTFGESHGLAGCIVDGVPPGIPTEADLQHDLDLRRRPGTSRYTTQRRE 61

QY 112 TDTKILSGTHGVTGTPIVIVPNTDQIGSDHREIANVYRPSHADATYDKYGVRAVQ 171

Db 62 PDQVKILSGVFGVTTGTSIGLLIENTDQRSQDSNAIKDVFRPGHADYTYEQYGLRDYR 121

QY 172 GGRSSARETAMRVAAGAAIAKYLAEKFGIEIRGCLTQMGGDI--PLEIKDW-----SQVE 174

Db 122 GGRSSARETAMRVAAGAAIAKYLAEKFGIEIRGCLTQMGGDI--PLEIKDW-----SQVE 174

QY 232 SNIVRCPPDEYAEKMDAIDR----VRVRGDSVGGVITTCVARNVPRGLSGSPVFDKLESEL 287

Db 175 QNPFPCPPD-----KIDALDELMRALKKEGDSIGAKVTIVVAGVPAGLGEVFDRLDADI 230

QY 288 AKAMLSIPASNGFEIGSGFAGTDLTGSEHNDEFYMDKAGSVRTRTNRSQVGGVGSINVEI 347

Db 231 AHALMSINAVKVGVEIGDGDVVALRGSQNRDEITKD-----GFQSNHAGGILGGISSGQQ 285

QY 348 VHFVKAFKTPPSIGVKQNTVSRERQNVELLARGRHDPVAPRAVPVSVESMAALVMDQLM 407

Db 286 IIAHMAKPTSSITVPGRTINRFGEEVEMITKGRHDPVCGIRAVPIAEAMLAIVLMDHLL 345

QY 408 AHVAQ 412

Db 346 RQRAQ 350

RESULT 15

A55510

Chorismate synthase (EC 4.2.3.5) - Vibrio anguillarum

C:Species: Vibrio anguillarum

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jun-2002

C:Accession: A55510

R:Chen, Q.; Actis, L.A.; Tolmasky, M.E.; Croasa, J.H.

J. Bacteriol. 176, 4226-4234, 1994

A:Title: Chromosome-mediated 2,3-dihydroxybenzoic acid is a precursor in the biosynthesis

A:Reference number: A55510; MUID:94292450; PMID:8021209

A:Accession: A55510

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-372 <CHE>

A:Cross-references: GB:L29562

C:Genetics:

A:Gene: aroC

C:Superfamily: chorismate synthase

C:Keywords: carbon-oxygen lyase; phosphorus-oxygen lyase

Query Match 38.6%; Score 850; DB 1; Length 372;
Best Local Similarity 46.0%; Pred. No. 2.3e-56;
Matches 172; Conservative 70; Mismatches 118; Indels 14; Gaps 4;

QY 47 LEVKASGNTFNGYFQVATYGESHGCGVGVCGPPIPLTEADLQVELDRRRPGQSRTIT 106

Db 8 IEEVWAGNSIQHFRVMTFGESHGIALGCIVDGCPGLEITEADLQIDLDRRRPGTSRYT 67

QY 107 STRKETDTCKILSGTHGVTGTPIVIVPNTDQIGSDHREIANVYRPSHADATYDFKY 166

Db 68 TORREADEVKILSGVFGVTKTGTSGILLIENTDQRTDYSIDKDKFRPGHADYTYHQY 127

QY 167 VRAVQGGRRSGRKTGVRAAGALPKILKLCGLEILSFVSKVHQVLPEDAVDYGSVT 226

Db 128 IRDYRGGRRSSARETAMRVAAGRIAKYLKQEGFVEIRAYLSQMGDVCII--DKVDW---- 181

QY 227 LEQIESNIVRCPPDEYAEKMDAIDRVRVRGDSVGGVITTCVARNVPRGLSGSPVFDKLE 286

Db 182 -NEIENNAFFCPDADKVAAPDQILDLKKEGDSIGAKIQVATNLVGLGEVFDRLDAD 240

QY 287 LAKAMLSIPASNGFEIGSGFAGTDLTGSEHNDEFYMDKAGSVRTRTNRSQVGGVGSIN 346

Db 241 IAHALMSINAVKVGVEIGDGDVQVQKGSQHRDPLTPN-----GFRSNHAGGILGGIS 295

QY 347 IVHFVKAFKTPPSIGVKQNTVSRERQNVELLARGRHDPVAPRAVPVSVESMAALVMD 406

Db 296 DIVASIALKPTSSITVPGDTITRTGEPQTITKGRHDPVCGIRAVPIAEAMLAIVLMD 355

QY 407 MAHVAQCCEMFALNT 420

Db 356 LRHRGQ--NFAVQT 367

Search completed: August 25, 2005, 19:31:39

Job time : 20 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 19:22:15 ; Search time 67 Seconds
(without alignments)
3278.835 Million cell updates/sec

Title: US-10-677-179-8

Perfect score: 2200

Sequence: 1 MTTVPKQQVAHSRARLAPR.....VAQCEMFALNALQEPVGSF 429

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1541	70.0	431	1 ARC2 LYCES	Q42885 lycopersico
2	1527	69.4	440	1 ARC1 LYCES	Q42884 lycopersico
3	1486.5	67.6	447	1 AROC_CORSE	P27793 corydalis s
4	1454	66.1	436	2 Q9ZL13	Q9ZL13 arabidopsis
5	1438.5	65.4	435	1 AROC_ATH	P57220 arabidopsis
6	1191.5	54.2	362	1 AROC_SYNY3	P23353 synechocyst
7	1186.5	53.9	362	1 AROC_ANASP	Q8YYP9 anabaena sp
8	1176.5	53.5	363	2 Q8DLM1	Q8DLM1 synechococc
9	1047	47.6	364	2 Q7V364	Q7V364 prochloroco
10	1038	47.2	363	2 Q7U9F0	Q7U9F0 synechococc
11	1023.5	46.5	362	1 AROC_PROWA	P46894 prochloroco
12	1007.5	45.8	362	2 Q7V4Y9	Q7V4Y9 prochloroco
13	951	43.2	395	1 AROC_SCHPO	Q74413 schizosacch
14	950.5	43.2	382	2 Q6AIF3	Q6AIP3 desulfotale
15	944	42.9	376	2 Q6CNY1	Q6CNY1 kluyveromyc
16	944	42.9	376	2 Q8FV38	Q6FV38 candida gla
17	943	42.9	376	2 Q7S5B9	Q7S5B9 ashbya gos
18	938	42.6	377	1 AROC_YEAST	P28777 saccharomyce
19	938	42.6	377	2 Q6BQN1	Q6BQN1 debaryomyce
20	928.5	42.2	380	2 Q72W01	Q72W01 leptospira
21	928.5	42.2	380	2 Q8F9N4	Q8F9N4 leptospira
22	924.5	42.0	375	2 Q6C8Q1	Q6C8Q1 yarrowia li
23	924	42.0	432	1 AROC_NEUCR	Q12640 neurospora
24	889.5	40.4	358	2 Q64PF7	Q64PF7 bacteroides
25	887	40.3	219	2 Q75W14	Q75W14 bryza sativ
26	874.5	39.8	358	2 Q8A602	Q8A602 bacteroides
27	867	39.4	367	2 Q649I0	Q649I0 uncultured
28	863	39.2	361	2 Q6LNN4	Q6LNN4 photobacter
29	862	39.2	360	1 AROC_YERPE	Q8D41 yersinia pe
30	862	39.2	361	2 Q668V5	Q668V5 yersinia ps
31	860	39.1	360	1 AROC_ECOLI	P12008 escherichia

32	858	39.0	361	1 AROC_VIBPA	Q87mm9 vibrio para
33	856	38.9	360	1 AROC_ECO57	P63610 escherichia
34	856	38.9	360	1 AROC_ECOL6	P63609 escherichia
35	853	38.8	361	1 AROC_VIBAN	P39198 vibrio angu
36	851	38.7	361	2 Q6D2M6	Q6D2M6 erwinia car
37	849	38.6	362	2 Q7UC49	Q7UC49 shigella fl
38	849	38.6	362	2 Q83Q05	Q83Q05 shigella fl
39	847.5	38.5	377	2 Q7MIT1	Q7MIT1 vibrio vuln
40	846.5	38.5	357	1 AROC_PASMU	P57840 pasteurella
41	845	38.4	361	1 AROC_VIBVU	Q8db42 vibrio vuln
42	845	38.4	365	1 AROC_METWA	Q8pw84 methanosarc
43	844	38.4	353	1 AROC_BUCAP	Q9zhe9 buchnera ap
44	843	38.3	360	1 AROC_SALTY	P57229 salmoneilla
45	843	38.3	365	1 AROC_METAC	Q8tt87 methanosarc

ALIGNMENTS

```

RESULT 1
ARC2 LYCES STANDARD; PRT; 431 AA.
AC Q42885;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Chorismate synthase 2, Chloroplast precursor (EC 4.2.3.5) (5-
  enolpyruvylshikimate-3-phosphate phospholyase 2).
GN Name=C52;
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. UC82B;
RX MEDLINE=94072719; PubMed=8251624;
RA Goerlach J., Schmid J., Amrhein N.;
RT "Differential expression of tomato (Lycopersicon esculentum L.) genes
  encoding shikimate pathway isoenzymes. II. Chorismate synthase.";
RL Plant Mol. Biol. 23:707-716(1993).
CC -!- FUNCTION: Catalyzes the last common step of the biosynthesis of
  aromatic amino acids, produced via the shikimic acid pathway.
CC -!- CATALYTIC ACTIVITY: 5-O-(1-carboxyvinyl)-3-phosphoshikimate =
  chorismate + phosphate.
CC -!- COFACTOR: Reduced flavin.
CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
  seventh step.
CC -!- SUBUNIT: Homotrimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Predominantly expressed in flowers and roots
  and, to a lesser extent, in stems, leaves, and cotyledons.
CC -!- TISSUE SPECIFICITY: Belongs to the chorismate synthase family.
CC -!- SIMILARITY: Belongs to the chorismate synthase family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; Z21791; CAA79854.1; -.
PIR; S40409; S40409.
InterPro; IPR000453; Chorismate synth.
Pfam; PF01264; Chorismate synt; 1.
ProDom; PD002941; Chorismate synth; 1.
TIGRFAMs; TIGR00033; aroc_1_1.
PROSITE; PS00787; CHORISMATE SYNTHASE_1; 1.
PROSITE; PS00788; CHORISMATE SYNTHASE_2; 1.
PROSITE; PS00789; CHORISMATE SYNTHASE_3; 1.
Aromatic amino acid biosynthesis; Lyase; Transit peptide.

```



```

[1] SEQUENCE FROM N.A., AND SEQUENCE OF 267-276 AND 307-320.
RX MEDLINE=92042037; PubMed=3718979;
RX Schaller A., Schmid J., Leibinger U., Amrhein N.;
RA "Molecular cloning and analysis of a cDNA coding for chorismate
RT synthase from the higher plant Corydalis sempervirens Pers.";
RL J. Biol. Chem. 266:21434-21438(1991).
RN [2]
RP SEQUENCE OF 306-320.
RX PubMed=2146922;
RX Schaller A., Windhofer V., Amrhein N.;
RA "Purification of chorismate synthase from a cell culture of the higher
RT plant Corydalis sempervirens Pers.";
RL Arch. Biochem. Biophys. 282:437-442(1990).
CC -1- FUNCTION: Catalyzes the last common step of the biosynthesis of
CC aromatic amino acids, produced via the shikimic acid pathway.
CC -1- CATALYTIC ACTIVITY: 5-O-(1-carboxyvinyl)-3-phosphoshikimate =
CC chorismate + phosphate.
CC -1- COFACTOR: Reduced flavin.
CC -1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC seventh step.
CC -1- SUBUNIT: Homotetramer.
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- PTM: The N-terminus is blocked.
CC -1- SIMILARITY: Belongs to the chorismate synthase family.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X60544; CAA43034.1; -.
DR PIR; A41197; A41197.
DR InterPro; IPR000453; Chorismate_synth.
DR Pfam; PF01264; Chorismate_synth; 1.
DR ProDom; PD002941; Chorismate_synth; 1.
DR TIGRFAMs; TIGR00033; aroC; 1.
DR PROSITE; PS00787; CHORISMATE_SYNTHASE_1; 1.
DR PROSITE; PS00788; CHORISMATE_SYNTHASE_2; 1.
DR PROSITE; PS00789; CHORISMATE_SYNTHASE_3; 1.
KW Aromatic amino acid biosynthesis; Chloroplast;
KW Direct protein sequencing; Lyase; transit peptide.
KW TRANSIT 1 57 Chloroplast (Potential).
FT CHAIN 58 447 Chorismate synthase.
SQ SEQUENCE 447 AA; 48100 MW; 7CACAFD93A6BAEP8 CRC64;
Query Match 67.68; Score 1486.5; DB 1; Length 447;
Best Local Similarity 70.4%; Pred. No. 3.9e-101;
Matches 280; Conservative 60; Mismatches 55; Indels 3; Gaps 2;
Qy 31 ASSLSRPAVHRCRTA--RLVKASGNTFGNVFQVATYGESHGGVGVCVIGSCPPIRLPTE 88
Db 37 SNGSQVLSVR-QTAPLKLVAQSSGSGKVFQVTTYGESHGGVGVCVIGDCPPRFISE 95
Qy 89 ADIQVELDRRRPGQSRITSTKRETDTCILSGTGHTGVTTPILVIVPNTDQIGSDHREI 148
Db 96 ADIQSLLDRRRPGQSRITTTKRETDTCIKYSGVADGFTTGSPIHISVENTDQRNDYSEM 155
Qy 149 ANVYRSHADATYDFKYGVRVAVCGGSGRKTGRVAAGALPKKILKLCGLLEILSFVS 208
Db 156 AKAYRSHADATYDFKYGVRSVQGGGRSSARETIGRVAAGALAKKILKAYAGTEVLAYS 215
Qy 209 KVHQVVLPEDAVDYGSVLTLEQIESNIVRCPDPEYAEKMIADIDRVVRGDSVGVITPVA 268
Db 216 QAHKVVLPGLVDHETLSLEQIESNIVRCPDSEYAEKMIADIDRVVRGDSVGVVTCIM 275
Qy 269 RNVPRLGSGSPVKLSEELAKAMLSIPASNGFEIGSGFAGTDTLTGSEHNDEFYMDKAGSV 328
Db 276 RNVPRLGSGSPVKLEAELAKACMSLPATGKFFGSGFTLTGSHNDEFYTDENGRI 335

```

Db 210 HHVLPPELVHDLNLTLEQIENNVRCNPPEYAEKMIADAVRTKGNVGGVTCIVRN 269
Qy 271 VPRGLGSPVFDKLESELAAMLSIPASNGFIFGSGFAGTDLTGSEHNDEFYMDKAGSVRT 330
Db 270 APRGLGTPVFDKLEAEALACACMSLPATKGFPGSGFAGTFLTGLEHNDEFYTDENGRI 329
Qy 331 RTNRSGGVQGGISNNVEIVHFVAFKPTPSIGVKQNTVSRQNVVELLARGHDPVAPRA 390
Db 330 RTNRSGGVQGGISNGEIIINRVAFKPTSTIGRKQNTVTRDKVETEMIARGHDPVAPRA 389
Qy 391 VPVVESMAALVMDQLMAHVAQCEMPALNTALQEPV 426
Db 390 VPVVEAMVALVLDQLMAQYAQCHLFPINPELQEP 425

RESULT 5
ID AROC ARATH STANDARD; PRT; 435 AA.
AC P57720;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Chorismate synthase, chloroplast precursor (EC 4.2.3.5) (5-enolpyruvylshikimate-3-phosphate phospholigase).
GN OrderedLocusNames=At1g48850; ORFNames=T24P22.3, F27K7.1;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712; DOI=10.1038/35048500;
RA Theologis A., Ecker J.R., Palm C.J., Federpiel N.A., Kaul S., White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y., Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W., Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K., Dunn P., Egtv P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y., Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L., Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E., Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B., Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P., Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A., Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I., Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D., Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M., Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D., Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M., Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana.";
RL Nature 408:816-820(2000).
RC NATURE 408:816-820(2000).
CC -!- FUNCTION! Catalyzes the last common step of the biosynthesis of aromatic amino acids, produced via the shikimic acid pathway (By similarity).
CC -!- CATALYTIC ACTIVITY: 5-O-(1-carboxyvinyl)-3-phosphoshikimate = chorismate + phosphate.
CC -!- COFACTOR: Reduced flavin (By similarity).
CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway; seventh step.
CC -!- SUBUNIT: Homotrimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Chloroplast (By similarity).
CC -!- SIMILARITY: Belongs to the chorismate synthase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AC084242; AAC50662.1; --
DR EMBL; AC084414; AAC29740.1; --
DR PIR; C96526; C96526.
DR InterPro; IPR000453; Chorismate_synth.
DR Pfam; PF01264; Chorismate synt; 1.
DR ProDom; PD002941; Chorismate synth; 1.
DR TIGRPFAM; TIGR00033; aroC; 1.
DR PROSITE; PS00787; CHORISMATE SYNTHASE_1; 1.
DR PROSITE; PS00788; CHORISMATE SYNTHASE_2; 1.
DR PROSITE; PS00789; CHORISMATE SYNTHASE_3; 1.
KW Aromatic amino acid biosynthesis; Chloroplast; Lyase; Transit peptide.
FT TRANSIT 1 49 Chloroplast (Potential).
FT CHAIN 50 435 Chorismate synthase.
SQ SEQUENCE 435 AA; 47161 MW; 822B264B6C5FFC94 CRC64;
Query Match 65.4%; Score 1438.5; DB 1; Length 435;
Best Local Similarity 67.4%; Pred. No. 1.3e-97;
Matches 267; Conservative 65; Mismatches 63; Indels 1; Gaps 1;
Qy 31 ASSSLRFAVHRCRTARLEVKASNTGNYFOVATYGBSHGGVCGVSGCPRIPLTEAD 90
Db 30 SSPAVQISLRTQTRKNPQIQATSSSYGTHPRVSTFGSHGGVCGCIIDGCPRIPLTESD 89
Qy 91 LQVELDRRRPQGRITSTRKETDTCKILSGTHGVTGTPILVIVPNTDQIGSDHRIAN 150
Db 90 LQFDLD-RRPQGRITTPRKETDTCRISSGVSEGMTTGTPIHVFPVNTDQGLDYSEMSV 148
Qy 151 VYRPSHADATYDPKYGVRVAVQGGRRSGRKTGVRGAAGLPKKILKUKGLEILSFVSKV 210
Db 149 AYRPSHADATYDMKYGVRSVQGGRRSARETIGRVAPGALAKKILKQFAGTEILAYVSQV 208
Qy 211 HQVLPEDAVDYGSVTLQIESNIVRCNPPEYAEKMIADAVRTKGNVGGVTCIVARN 270
Db 209 HHVLPPELVHDLNLTLEQIENNVRCNPPEYAEKMIADAVRTKGNVGGVTCIVRN 269
Qy 271 VPRGLGSPVFDKLESELAAMLSIPASNGFIFGSGFAGTDLTGSEHNDEFYMDKAGSVRT 330
Db 269 APRGLGTPVFDKLEAEALACACMSLPATKGFPGSGFAGTFLTGLEHNDEFYTDENGRI 328
Qy 331 RTNRSGGVQGGISNNVEIVHFVAFKPTPSIGVKQNTVSRQNVVELLARGHDPVAPRA 390
Db 329 RTNRSGGVQGGISNGEIIINRVAFKPTSTIGRKQNTVTRDKVETEMIARGHDPVAPRA 388
Qy 391 VPVVESMAALVMDQLMAHVAQCEMPALNTALQEPV 426
Db 389 VPVVEAMVALVLDQLMAQYAQCHLFPINPELQEP 424

RESULT 6
ID AROC SYN3 STANDARD; PRT; 362 AA.
AC P23353;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Chorismate synthase (EC 4.2.3.5) (5-enolpyruvylshikimate-3-phosphate phospholigase).
GN Name=aroC; OrderedLocusNames=sl11747;
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94086566; PubMed=7505271;
RA Schmidt J., Bubunenko M., Subramanian A.R.;
RT "A novel operon organization involving the genes for chorismate synthase (aromatic biosynthesis pathway) and ribosomal GTPase center proteins (L11, L1, L10, L12: rplKALJ) in cyanobacterium Synechocystis PCC 6803.";
RT PCC 6803.";
RL J. Biol. Chem. 268:27447-27457(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;

RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugiura M., Sasamoto S., Kimura T.,
RA Hosouchi T., Hataeno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT *Synechocystis* sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
RN [3]
RP SEQUENCE OF 1-138 FROM N.A.
RX MEDLINE=91002677; PubMed=2119815; DOI=10.1016/0167-4781(90)90142-O;
RA Sibold C., Subramanian A.R.;
RT "Cloning and characterization of the genes for ribosomal proteins L10
RT and L12 from *Synechocystis* sp. PCC 6803: comparison of gene clustering
RT pattern and protein sequence homology between cyanobacteria and
RT chloroplasts.";
RL Biochim. Biophys. Acta 1050:61-68(1990).
CC -!- CATALYTIC ACTIVITY: 5-O-(1-carboxyvinyl)-3-phosphoshikimate =
CC chorismate + phosphate.
CC -!- COFACTOR: Reduced flavin (By similarity).
CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC seventh step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SIMILARITY: Belongs to the chorismate synthase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X67516; CAA47855.1; -
CC EMBL; D90906; BAA17415.1; -
CC EMBL; X53178; CAA37319.1; -
CC PIR; A49316; A49316.
CC HAMAP; MF_00300; -; 1.
CC InterPro; IPR000453; Chorismate_synth.
CC Pfam; PF01264; Chorismate_synth; 1.
CC ProDom; PD002941; Chorismate_synth; 1.
CC TIGRFAMs; TIGR00033; aroC; 1.
CC PROSITE; PS00787; CHORISMATE SYNTHASE 1; 1.
CC PROSITE; PS00788; CHORISMATE SYNTHASE 2; 1.
CC PROSITE; PS00789; CHORISMATE SYNTHASE 3; 1.
CC Aromatic amino acid biosynthesis; Complete proteome; Lyase.
CC SEQUENCE 362 AA; 39287 MW; 9709B0F409168AB8 CRC64;

CC Query Match 54.2%; Score 1191.5; DB 1; Length 362;
CC Best Local Similarity 62.2%; Pred. No. 1.6e-79;
CC Matches 225; Conservative 57; Mismatches 77; Indels 3; Gaps 1;

QY 53 GNTFGNYFOVATYGESHGCGVCGVIGSCPPRIPLTEADLQVELDRRRPGQSRITSTRKET 112
Db 2 GNTFGSLPRITTFGSHGGGVGVIIDGCPPLRISPEEIQVLDLDRRRPGQSKITTPRKEA 61

QY 113 DTCKILSGTHEGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGVRVQ 172
Db 62 DQCEILSGVFEKTLGTPTAILVRNKDARSODYNEMAVKYPFSHADATYEAQYGINWQ 121

QY 173 GGRSSGRKTVGRAAGALPKKILKLGLEILSFYSKVHVVLPDAVDYGSVTLQEQIES 232
Db 122 GGRSSARETIGRVAAGAIKAIQFNGVEIVAYVKS IQDI---RATVDSNTVTLQEVES 178

QY 233 NIVRCPPDEYAEKMDAIDRVVRGDSVGGVITCVARNVPRGLGSPVDFDKLESLAKAML 292
Db 179 NIVRCPPDECAEKMIERIDQVLQRKDSIGGVVECAIRNAPKGLGEPVDFDKLEADLAKAM 238

QY 293 SIPASNGEIGSGFAGTDLTGSEHNDEFYMDXAGSVRTTRNSGGVQGSINVEIVHFKV 352
Db 239 SLPATKGFEGSGFAGTLLTGSHNDEYYLDEAGEWRTRNRSRGGVQGSINSGEPIIMRI 298

QY 353 AFKPTPSIGVKQNTVSRERQNVELLARGHDPVAPRAVPVVSMAALVLMQDMAHVAQ 412
Db 299 AFKPTATIGQSKQNTVSGNIGETTLAARGHDPVCLPRAVPVVEAALVLCDDLRLRFOAQ 358

QY 413 CE 414
Db 359 CK 360

RESULT 7

AROC ANASP
ID AROC ANASP STANDARD; PRT; 362 AA.
AC Q8YFP9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Chorismate synthase (EC 4.2.3.5) (5-enolpyruvylshikimate-3-phosphate
DE phosphohylase).
GN Name=aroC; OrderedLocuNames=all0797;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriquchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohza M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpou S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium *Anabaena* sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
CC -!- CATALYTIC ACTIVITY: 5-O-(1-carboxyvinyl)-3-phosphoshikimate =
CC chorismate + phosphate.
CC -!- COFACTOR: Reduced flavin (By similarity).
CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC seventh step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SIMILARITY: Belongs to the chorismate synthase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP003583; BAB72754.1; -
CC PIR; AC1906; AC1906.
CC HAMAP; MF_00300; -; 1.
CC InterPro; IPR000453; Chorismate_synth.
CC Pfam; PF01264; Chorismate_synth; 1.
CC ProDom; PD002941; Chorismate_synth; 1.
CC TIGRFAMs; TIGR00033; aroC; 1.
CC PROSITE; PS00787; CHORISMATE SYNTHASE 1; 1.
CC PROSITE; PS00788; CHORISMATE SYNTHASE 2; 1.
CC PROSITE; PS00789; CHORISMATE SYNTHASE 3; 1.
CC Aromatic amino acid biosynthesis; Complete proteome; Lyase.
CC SEQUENCE 362 AA; 39339 MW; 4DD32588F01C22FC CRC64;

CC Query Match 53.9%; Score 1186.5; DB 1; Length 362;
CC Best Local Similarity 60.6%; Pred. No. 3.8e-79;
CC Matches 220; Conservative 66; Mismatches 74; Indels 3; Gaps 1;

QY 53 GNTFGNYFOVATYGESHGCGVCGVIGSCPPRIPLTEADLQVELDRRRPGQSRITSTRKET 112
Db 2 GNTFGSLPRITTFGSHGGGVGVIIDGCPPLRISPEEIQVLDLDRRRPGQSKITTPRKEA 61

QY 113 DTCKILSGTHEGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGVRVQ 172
Db 62 DTCBILSGVFEKTLGTPTAILVRNKDTRPDYDQYDQYDQYDQYDQYDQYDQYDQYDQY 121

Query Match 47.6%; Score 1047; DB 2; Length 364;
Best Local Similarity 54.9%; Pred. No. 7.2e-69;
Matches 200; Conservative 64; Mismatches 96; Indels 4; Gaps 2;

QY 53 GNTFGNYFOVATYGGSHGGVGVISGCPPIPLTEADLOVELDRRRPQSQSRTSTRKET 112
DB 2 GSSFGKQFRVTTFGSHGGVGVILDCPPKLNIDLIQNELDRRRPQSQKITTPRNEB 61
QY 113 DTKILSGTHGVTCTPILVIVPNTDQIGSHREIANVVRPASHADATYDFKYGVRVAVG 172
DB 62 DKLEILSLGKEGTLGTPIAMVRNKKDQRPDYSNLEQVFRSHADGTHLTKIGIQAGSG 121
QY 173 GGRSSGRKTVGRVAAGALPKKILKLCGLLEILSFVSKVHQVLPEDAVDYGVSVTLEQIES 232
DB 122 GGRASARETIGRVAAGAIKALLKLNLFTEILLSWVKRIHDI--DSQVNNKNTLSKIDS 178
QY 233 NVIRCPDPPYAEKMDAIDRVRVGRSGVGTCTVARNVPRGLGSPVDFKLESLAKAML 292
DB 179 NVIRCPDDKVAAMKRIKELQODGDSGCVIECLVKNVPSGLGMPVDFKLEADLAKALM 238
QY 293 SIPASNGFEIGSGFAGTDLTSGSHNDDEFY-MDKAGSVTRTRNRSQGVGSGISNVEIVHFK 351
DB 239 SIPATKGFEGSGFLGTLYRGSHNDSTFESDIIKLTISNNSGTOGGISNGENIENK 298
QY 352 VAFKPTPSIGVQNTVSRERQNVNELLARGHDPVAPRAVPVSVESMAALVMDQLMAHVA 411
DB 299 IAFKPTATIGKQKTVNADGKELMKAKGRHDPVLPRAVPVDSNVALVLDHLLHQA 358
QY 412 QCEN 415
DB 359 QCSL 362

RESULT 10

QYU9FO PRELIMINARY; PRT; 363 AA.

ID QYU9FO
AC QYU9FO
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-WAR-2004 (Tremblrel. 26, Last annotation update)
DE Chorismate synthase (EC 4.2.3.5).
GN Name=aroC; OrderedLocusNames=SYNW0308;
OS Synchococcus sp. (strain WH8102).
OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.
OX NCBI_TaxID=84588;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825697; PubMed=12917641; DOI=10.1038/nature01943;
RA Palenik B., Brahamsha B., Larimer F.W., Land M.L., Hauser L.,
Chain P., Lamerdin J.E., Regala W., Allen E.B., McCarren J.,
Paulsen I.T., Dufresne A., Partensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synchococcus."
RL Nature 424:1037-1042 (2003).
CC -!- CATALYTIC ACTIVITY: 5-O- (1-carboxyvinyl)-3-phosphoshikimate =
chorismate + phosphate.
CC -!- COFACTOR: Reduced flavin (By similarity).
CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
seventh step.
CC -!- SIMILARITY: Belongs to the chorismate synthase family.
DR EMBL; BX569689; CAB06823.1; -.
DR HSP; P28777; I152.
DR GO; GO:0004107; F:chorismate synthase activity; IEA.
DR GO; GO:0016029; F:lyase activity; IEA.
DR GO; GO:0009073; P:aromatic amino acid family biosynthesis; IEA.
DR InterPro; IPR000453; Chorismate synth.
DR Pfam; PF01264; Chorismate synth; 1.
DR ProDom; PD002941; Chorismate synth; 1.
DR TIGRFAMs; TIGR00033; aroC; 1.
DR PROSITE; PS00787; CHORISMATE SYNTHASE 1; 1.
DR PROSITE; PS00788; CHORISMATE SYNTHASE 2; 1.
DR PROSITE; PS00789; CHORISMATE SYNTHASE 3; 1.
KW Amino-acid biosynthesis; Aromatic amino acid biosynthesis;
Complete proteome; Lyase.

SQ SEQUENCE 363 AA; 38266 MW; CDP87180AF34BF70 CRC64;
Query Match 47.2%; Score 1038; DB 2; Length 363;
Best Local Similarity 54.8%; Pred. No. 3.3e-68;
Matches 200; Conservative 65; Mismatches 96; Indels 4; Gaps 2;

QY 53 GNTFGNYFOVATYGGSHGGVGVISGCPPIPLTEADLOVELDRRRPQSQSRTSTRKET 112
DB 2 GSSFGDLFRISTFGSHGGVGVIVGCPPLDLDDVDAIOAELDRRRPQSQSHITTPRKEA 61
QY 113 DTKILSGTHGVTCTPILVIVPNTDQIGSHREIANVVRPASHADATYDFKYGVRVAVG 171
DB 62 DQVELSLGLLDGQITLTGTPIAMLVNKKDQRPDYSNLEQVFRSHADATYDFKYGVRVAVG 121
QY 172 GGRSSGRKTVGRVAAGALPKKILKLCGLLEILSFVSKVHQVLPEDAVDYGVSVTLEQIE 231
DB 122 GGRASARETIGRVAAGAIKALLKLNLFTEILLSWVKRIHTI--EAVNVDPAVTLDAIE 178
QY 232 SNIVRCPDPPYAEKMDAIDRVRVGRSGVGTCTVARNVPRGLGSPVDFKLESLAKAM 291
DB 179 SNIVRCPDPAQAQVVERIEAIGRDSGCVIECVNRPAPGLGMPVDFKLEADLAKAV 238
QY 292 LSIPASNGFEIGSGFAGTDLTSGSHNDDEFYMDKAGSVTRTRNRSQGVGSGISNVEIVHFK 351
DB 239 MSLPATKGFEGSGFSGTLKSGSHNDAFVPTDDGRLQTATNNSGGIOGGISNGEPIVIR 298
QY 352 VAFKPTPSIGVQNTVSRERQNVNELLARGHDPVAPRAVPVSVESMAALVMDQLMAHVA 411
DB 299 VAFKPTATIRKEQQTIDSGKATTLAGKGRHDPVLPRAVPVMEAMVALVLDHLLRQOG 358
QY 412 QCENF 416
DB 359 QCSLM 363

RESULT 11

AROC PROMA
ID AROC PROMA STANDARD; PRT; 362 AA.

AC P46894;
DT 01-NOV-1995 (Rel. 32, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Chorismate synthase (EC 4.2.3.5) (5-enolpyruvylshikimate-3-phosphate
phosphorylase).
GN Name=aroC; OrderedLocusNames=Pro0253;
OS Prochlorococcus marinus.
OC Bacteria; Cyanobacteria; Prochlorales; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1219;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SARG / CCM 1375 / SS120;
RX MEDLINE=22810154; PubMed=12917486; DOI=10.1073/pnas.1733211100;
RA Dufresne A., Sallanoubat M., Partensky F., Artiguenave F., Axmann I.M.,
Barbe V., Duprat S., Galperin M.V., Koonin E.V., Le Gall F.,
Makarova K.S., Ostrowski M., Oztas S., Robert C., Rogozin I.B.,
Scanlan D.J., Tandeau de Marsac N., Weissbach J., Winzler P.,
Wolf Y.I., Hess W.R.;
RT "Genome sequence of the cyanobacterium Prochlorococcus marinus SS120,
a nearly minimal oxyphototrophic genome."
RL Proc. Natl. Acad. Sci. U.S.A. 100:10020-10025 (2003).
RN [2]
RP SEQUENCE OF 1-52 FROM N.A.
RC STRAIN=SARG / CCM 1375 / SS120;
RX MEDLINE=95284368; PubMed=7766900;
RA Hess W.R., Weihe A., Loiseux-De Goer S., Partensky F., Vault D.;
RT "Characterization of the single psbA gene of Prochlorococcus marinus
CCMP 1375 (Prochlorophyta)."
RL Plant Mol. Biol. 27:1189-1196 (1995).
CC -!- CATALYTIC ACTIVITY: 5-O- (1-carboxyvinyl)-3-phosphoshikimate =
chorismate + phosphate.
CC -!- COFACTOR: Reduced flavin (By similarity).
CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;

```
CC seventh step.
CC -!- SURUNIT: Homotetramer (By similarity).
CC -!- SIMILARITY: Belongs to the chorismate synthase family.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE017161; AAP92929.1; -.
DR EMBL; Z49201; CAA89063.1; -.
DR HAMAP; MF_00300; -, 1
DR InterPro; IPR000453; Chorismate_synth.
DR Pfam; PF01264; Chorismate_synth; 1.
DR ProDom; PD002941; Chorismate_synth; 1.
DR TIGRFAMs; TIGR00033; aroC; 1.
DR PROSITE; PS00787; CHORISMATE_SYNTHASE_1; 1.
DR PROSITE; PS00788; CHORISMATE_SYNTHASE_2; 1.
DR PROSITE; PS00789; CHORISMATE_SYNTHASE_3; 1.
KW Aromatic amino acid biosynthesis; Complete proteome; Lyase.
FT CONFLICT 49 52 PQQS -> RQQR (in Ref. 2).
SQ SEQUENCE 362 AA; 39429 MW; 4D761D46A694506C CRC64;

Query Match 46.5%; Score 1023.5; DB 1; Length 362;
Best Local Similarity 52.5%; Pred. No. 3.8e-67;
Matches 191; Conservative 69; Mismatches 101; Indels 3; Gaps 1;

QY 53 GNTFGNYFOVATYGESHGCGVCGVIGCPPIPLTEADLQVELDRRRPGQSRTITSRKET 112
DB 2 GSSFGDLFRVSTFGESHGSGVGIVGECPPRLDLEKIQEELDRRRPGQSKITPRKEL 61

QY 113 DTCILSGTHEGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDPKYGVRAVQG 172
DB 62 DQVEILSGIANRETIGTPIAMVRNDRQPSDYKEMKIFRPSHADGTYHLKYGIQAPSG 121

QY 173 GGRSSGRKTGVRVAGALPKKILKLCGLEILSFVSKVQHVLPEDAVDYGSVTLEQIES 232
DB 122 GGRASARETIQKVAAGAKQLQKQVNTQILAWVKRIHNI---EAEIDINTIGFADIES 178

QY 233 NIVRCPDPEYAEKMDAIDRVRVRGDSVGGVITCVARNVPRGLSPVPDKLESELAKAML 292
DB 179 NIVRCPNQDVAKMLQRIEESRDGSCGGLIECVRVNVPAGLGNPVPDKLEADLSKALM 238

QY 293 SIPASNGFEIGSGFAGTDLTSEHNDEFYMDKAGSVRTRNRSRGVQGGISNVEIVHFKV 352
DB 239 SLPATKGPVSGVSGFRGTFLKGSSEHNDAFIAGDKNRLRTATNNSGGIQGGISNGSPILRV 298

QY 353 AFKPTPSIGVKQNTVSRERQNVELLARGHDPVAPRAVPVVSMAALVMDQLMAHVAQ 412
DB 299 GFKPATIRKQOQTIDSEKQITLASKGRHDPVCLPRAVPVMEAMVSVTLADHLLRQGG 358

QY 413 CEMF 416
DB 359 CSLW 362

PRELIMINARY; PRT; 362 AA.

RESULT 12
Q7V4Y9
ID Q7V4Y9 PRELIMINARY; PRT; 362 AA.
AC Q7V4Y9
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Chorismate synthase (EC 4.2.3.4).
GN Name=aroC; OrderedLocNames=PMT1795;
OS Prochlorococcus marinus (strain MIT 9313).
OC Bacteria; Cyanobacteria; Prochlorales; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=74547;
RN [1]
```

```
RP SEQUENCE FROM N. A.
RX MDLINE=22825698; PubMed=12917642; DOI=10.1038/nature01947;
RA Rocap G., Larimer F.W., Lamerdin J.E., Malfatti S., Chain P.,
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA Johnson Z.I., Land M.L., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinner E.R., Chisholm S.W.;
RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic
RT niche differentiation.";
RL Nature 424:1042-1047(2003).
CC -!- CATALYTIC ACTIVITY: 5-O- (1-carboxyvinyl) -3-phosphoshikimate =
CC chorismate + phosphate.
CC -!- COFACTOR: Reduced flavin (By similarity).
CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC seventh step.
CC -!- SIMILARITY: Belongs to the chorismate synthase family.
CC EMBL; BX572100; CAE21970.1; -.
CC HSP; P28777; IR52.
CC GO; GO:0003856; F:3-dehydroquinate synthase activity; IEA.
CC GO; GO:0004107; F:chorismate synthase activity; IEA.
CC GO; GO:0016829; F:lyase activity; IEA.
CC GO; GO:0016829; F:lyase activity; IEA.
CC GO; GO:0009073; P:aromatic amino acid family biosynthesis; IEA.
CC InterPro; IPR000453; Chorismate_synth.
CC Pfam; PF01264; Chorismate_synth; 1.
CC ProDom; PD002941; Chorismate_synth; 1.
CC TIGRFAMs; TIGR00033; aroC; 1.
CC PROSITE; PS00787; CHORISMATE_SYNTHASE_1; 1.
CC PROSITE; PS00788; CHORISMATE_SYNTHASE_2; 1.
CC PROSITE; PS00789; CHORISMATE_SYNTHASE_3; 1.
KW Amino-acid biosynthesis; Aromatic amino acid biosynthesis;
KW Complete proteome; Lyase.
SQ SEQUENCE 362 AA; 38714 MW; 435DEA6777BB86D3 CRC64;

Query Match 45.8%; Score 1007.5; DB 2; Length 362;
Best Local Similarity 51.6%; Pred. No. 5.8e-66;
Matches 188; Conservative 69; Mismatches 104; Indels 3; Gaps 1;

QY 53 GNTFGNYFOVATYGESHGCGVCGVIGCPPIPLTEADLQVELDRRRPGQSRTITSRKET 112
DB 2 GSSFGDLFRVSTFGESHGSGVGIVGECPPRLDLEKIQEELDRRRPGQSKITPRKEE 61

QY 113 DTCILSGTHEGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDPKYGVRAVQG 172
DB 62 DQVEILSGLLNNTLTGPIAMVRNDRQHPGDKYKEMVAFRPSHADATYQAKYGIQARSQ 121

QY 173 GGRSSGRKTGVRVAGALPKKILKLCGLEILSFVSKVQHVLPEDAVDYGSVTLEQIES 232
DB 122 GGRASARETIQKVAAGAKQLLTKAHTEVLAWVKRIHTL---EAFINAQDVSIDDVEA 178

QY 233 NIVRCPDPEYAEKMDAIDRVRVRGDSVGGVITCVARNVPRGLSPVPDKLESELAKAML 292
DB 179 NIVRCPNQVMAQNVRIEIAISREGUSCGVIECVRVNAPGLMPVPDKLEADLAKAVM 238

QY 293 SIPASNGFEIGSGFAGTDLTSGSEHNDEFYMDKAGSVRTRNRSRGVQGGISNVEIVHFKV 352
DB 239 SLPAKSGFEIGSGFGGTLLKGSSEHNDALFSPNDRLRTATNNSGGIQGGITNGESIVRV 298

QY 353 AFKPTPSIGVKQNTVSRERQNVELLARGHDPVAPRAVPVVSMAALVMDQLMAHVAQ 412
DB 299 AFKPTATIRKQOQTIDAGNTTTLAKGRHDPVCLPRAVPVIVMEAMVSLVLADHLLRQGG 358

QY 413 CEMF 416
DB 359 CSLW 362

RESULT 13
AROC_SCHPO
ID AROC_SCHPO STANDARD; PRT; 395 AA.
AC Q74413; Q9Y7Q4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
```

Chorismate synthase (EC 4.2.3.5) (5-enolpyruvylshikimate-3-phosphate phospholyase).

DE ORFNames=SPCC1223.14, SPCC297.01;
 GN Schizosaccharomyces pombe (Pisision yeast).
 OS Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
 RA Wood V., Gilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Stevens S., Warren T.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Furnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardet C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -!- CATALYTIC ACTIVITY: 5-O-(1-carboxyvinyl)-3-phosphoshikimate =
 CC chorismate + phosphate.
 CC -!- COFACTOR: Reduced flavin (By similarity).
 CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
 CC seventh step.
 CC -!- SUBUNIT: Homotetramer (By similarity).
 CC -!- SIMILARITY: Belongs to the chorismate synthase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by not-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AL031579; CAA20883.1; -;
 DR EMBL; AL049609; CAB40781.1; -;
 DR FIR; T41268; T41268
 DR GeneDB_Spombe; SPCC1223.14; -;
 DR InterPro; IPR000453; Chorismate synth.
 DR Pfam; PF01264; Chorismate_synth; 1.
 DR ProDom; PD002941; Chorismate synth; 1.
 DR TIGRFAMs; TIGR00033; arcc; 1;
 DR PROSITE; PS00787; CHORISMATE SYNTHASE 1; 1.
 DR PROSITE; PS00788; CHORISMATE SYNTHASE 2; 1.
 DR PROSITE; PS00789; CHORISMATE SYNTHASE 3; 1.
 KW Aromatic amino acid biosynthesis; lysase.
 SQ SEQUENCE 395 AA; 42325 MW; B551505160696DE5 CRC64;
 Query Match 43.2%; Score 951; DB 1; Length 395;
 Best Local Similarity 51.3%; Pred. No. 9.4e-62;
 Matches 194; Conservative 57; Mismatches 103; Indels 24; Gaps 4;
 QY 54 NTFGNVQVATYGESHGGVGCVISGCPPIPLTEADLQVLDLRRRRPGQSRTSTRKETD 113

Db 2 SSFGTLFKVTVTGESHCCKSVGCVGCPFGNNLTESDVQVOLTTRRRPGQSNLTTPRNEKD 61
 QY 114 TKLILSGTHEGVTGTTPILVVPNTDQIGSDHREITANVVRPSHADATYDFKYGVRVQGG 173
 Db 62 KVQIQSGTFEGVTGTPTGLMLVNDQKPHDYSMDNDNPRPSHADYTYLEKYGVKASGG 121
 QY 174 GRSSGRKTVGRVAAGALPKILKLCGLLEILSFVSKVHQVVL-----PED--- 218
 Db 122 GRSSARETIGRVAAGIAEKYLLLEAYGVEIVAFVSSVGKIALPLHETASSAILDEDDTF 181
 QY 219 ----AVDY-----GSVTLQIESNIVRCPDPPEYAEKMDAIDRVRVRGDSVGVGVTVCARN 270
 Db 182 ESPITAELKFLNKITREEDVKTTVRCFPAATAAKMAERITRARDNHSIGGTVTCVIRN 241
 QY 271 VPRGIGSPVFKLESELAKAMLSIPASNGFEGISGAGTDLTGSEHDEFFYMD-KAGSVR 329
 Db 242 VPTGLGEFCFKLEAKLAHAMWSIPATKSPFISGREGCKVAGSKHNDLFYRNADTGKLG 301
 QY 330 TRTRNSGGVQGGISNVEIVHFVKAPKTPSPISGVKONTVSRERQNVELLARGHDPVAPR 389
 Db 302 TLTNNSGGVQGGISNGENVYTFGKSPATIGVEQSTSRYDGSQGVLAAGKGRHDPVAPR 361
 QY 390 APVVVESMAALVLMQDL 407
 Db 362 AIPIVEAMALVWMDAVM 379
 ID Q6AIP3 PRELIMINARY; PRT; 382 AA.
 AC Q6AIP3;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DE Probable chorismate synthase.
 GN OrderedLocustNames=DP3058;
 OS Desulfotalea psychrophila.
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacteriales;
 OC Desulfobubaceae; Desulfotalea.
 OX NCBI_TaxID=84980;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LSV54 / DSM 12343;
 RX PubMed=15305914;
 RA Rabus R., Ruepp A., Frickey T., Rattei T., Partmann B., Stark M.,
 RA Bauer M., Zibat A., Lombardot T., Becker I., Amann J., Gellner K.,
 RA Teeling H., Leuschner W.D., Gloeckner P.-O., Lupas A.N., Amann R.,
 RA Klenk H.-P.;
 RT "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium
 RT from permanently cold Arctic sediments.";
 RL Environ. Microbiol. 6:887-902(2004).
 CC -!- CATALYTIC ACTIVITY: 5-O-(1-carboxyvinyl)-3-phosphoshikimate =
 CC chorismate + phosphate.
 CC -!- COFACTOR: Reduced flavin (By similarity).
 CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
 CC seventh step.
 CC -!- SIMILARITY: Belongs to the chorismate synthase family.
 DR EMBL; CR522870; CAG37787.1; -;
 DR GO; GO:0004107; P:chorismate synthase activity; IEA.
 DR GO; GO:0009073; P:aromatic amino acid family biosynthesis; IEA.
 DR InterPro; IPR000453; Chorismate synth.
 DR Pfam; PF01264; Chorismate_synth; 1.
 DR ProDom; PD002941; Chorismate synth; 1.
 DR TIGRFAMs; TIGR00033; arcc; 1.
 DR PROSITE; PS00787; CHORISMATE SYNTHASE 1; 1.
 DR PROSITE; PS00788; CHORISMATE SYNTHASE 2; 1.
 DR PROSITE; PS00789; CHORISMATE SYNTHASE 3; 1.
 KW Amino-acid biosynthesis; Aromatic amino acid biosynthesis;
 KW Complete proteome; lysase.
 SQ SEQUENCE 382 AA; 41321 MW; 5A93D8AF1802F4A CRC64;
 Query Match 43.2%; Score 950.5; DB 2; Length 382;


```
Best Local Similarity 51.8%; Pred. No. 9,8e-62;
Matches 188; Conservative 65; Mismatches 107; Indels 3; Gaps 2;

QY 54 NTFGNFYQVATYGESHGGVGVCGVISCPCPRIPLEADLQVELDRRRPQGSRIITSTRKETD 113
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 23 SSFGTLXKYSTYGESHCKGKGVVWIDGCPGLESEADIQQLDRRRPQGVNLTTRKEAD 82
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 114 TCKILSGTHGVTGTPILVIVPNTDQIGSDHREIANVYRSHADATYDFKYGVRVQGG 173
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 83 QVMILSGTENGKTLGTPIALHVNRRDQRPDGYDISLPRPSHADFTYQMKYGRASSGG 142
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 174 GRSSGRKTVGRVAAGALPKKILKCGLEILSFVSKVHVLPEDAVDYGVSVTLEQIESN 233
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 143 GRASARETIGTVAAGAVKMLKKYGIIEIVAWVESVGP1--QASGINPDTISREVDKQ 200
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 234 IVRCPDPEYAEKMDAIDRVRVRGDSVGVITCVARNVPRGLSGVPFKLSELAAMLS 293
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 201 ISRCPDQITAAEMERKFIETELKERGDSGTGIVTCCIRNSFVGLGEPIYEKLEAKLAQMLA 260
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 294 IPASNGFPIGSGFAGTDLTGSEHNDEFYMDKAGSVRTNRSVGVOGGISNVEIVHFKA 353
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 261 IPASGPIGSGFAGSROLGSEHNDLFTW-KDNRLGTSTNNSGGIQQGSGISNGEPIILFRVA 319
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 354 FKPTPSIGVKQNTVSRERQNVELLARGRHDPCVAPRAVPVVSMAALVMDQMAHVAQC 413
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 320 FKPTATISLPQETATFTGKTETMQAKGRHDPCVSRAP1P1VESMAALVLLDMWRQESRK 379
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 414 EMF 416
   |
Db 380 SPF 382

RESULT 15
Q6CNY1 PRELIMINARY; PRT; 376 AA.
AC Q6CNY1;
DC 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Kluyveromyces lactis strain NRRL Y-1140 chromosome E of strain NRRL Y-
DE 1140 of Kluyveromyces lactis.
GN ORFNames=KLI1A0E09031G;
OS Kluyveromyces lactis NRRL Y-1140.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=284590;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RG Genolevures;
RA Dufon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neveuglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boirame A., Boyer J., Catholico L., Confaniolieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekaja P., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zenlou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Galliard C., Weissenbach J.,
RA Winkler P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: 5-O-(1-carboxyvinyl)-3-phosphoshikimate =
   chorismate + phosphate.
CC
```

CC -1- COFACTOR: Reduced flavin (By similarity).
CC -1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC seventh step.
CC -1- SIMILARITY: Belongs to the chorismate synthase family.
DR EMBL; CR382125; CAG9445.1; -.
DR GO; GO:0004107; P:chorismate synthase activity; IEA.
DR GO; GO:0009073; P:aromatic amino acid family biosynthesis; IEA.
DR InterPro; IPR000453; Chorismate_synth.
DR Pfam; PF01284; Chorismate_synth; 1.
DR ProDom; PD002941; Chorismate_synth; 1.
DR TIGRFAMs; TIGR00033; aroc; 1.
DR PROSITE; PS00787; CHORISMATE SYNTHASE 1; 1.
DR PROSITE; PS00788; CHORISMATE SYNTHASE 2; 1.
DR PROSITE; PS00789; CHORISMATE SYNTHASE 3; 1.
KW Amino-acid biosynthesis; Aromatic amino acid biosynthesis; Lyase.
SQ SEQUENCE 376 AA; 40348 MW; D4CB8B2B4DF170A9 CRC64;

Query Match 42.9%; Score 944; DB 2; Length 376;
Best Local Similarity 51.9%; Pred. No. 2.9e-61;
Matches 188; Conservative 62; Mismatches 104; Indels 8; Gaps 3;
QY 54 NTFGNFYQVATYGESHGGVGVCGVISCPCPRIPLEADLQVELDRRRPQGSRIITSTRKETD 113
Db 2 STFGQIFRVATYGESHCCKSVGCIVDGVPPGMSLTDDIQQLTRRRPQSKLSTPRNEKD 61
QY 114 TCKILSGTHGVTGTPILVIVPNTDQIGSDHREIANVYRSHADATYDFKYGVRVQGG 173
Db 62 RVEIQSGTEFGKTLGTPIGMLVRNEDQRPHDYSMDNYPSPSHADFTYSSKYGKAKSSGG 121
QY 174 GRSSGRKTVGRVAAGALPKKILKCGLEILSFVSKVHVLPEDAVD-----YGSVTL 227
Db 122 GRSSARETIGRVAAGAIAEKFLQLSNVEIVAFVTOIGAVKMRDPODPKPFQVNLITR 181
QY 228 EQIE-SNIVRCPPDEYAEKMDAIDRVRVRGDSVGVITCVARNVPRGLSGVPFKLSE 286
Db 182 EKVDAAAGPIRCPDSSVADAMVKGIERYRGEKDSIGGVVTCVIRNLPAGLGEPCPKLEAL 241
QY 287 LAKAMLSIPASNGFPIGSGFAGTDLTGSEHNDEFYMDK-AGSVTRTRNRSVGVOGGISNV 345
Db 242 LAHAWLSIPASKGFEIGSGFDGVCPLGSKHNDPFYFDEDTGRUKTKTNNSGGIQQGINSNG 301
QY 346 EIVHFKVAFKPTPSIGVKQNTVSRERQNVELLARGRHDPCVAPRAVPVVSMAALVMDQ 405
Db 302 ENIYFSVPFKSAATISQEQATATYDGKSGILAAKGRHDPSTVTPRATPIVEAMAAVLADA 361
QY 406 LM 407
Db 362 VL 363

Search completed: August 25, 2005, 19:31:17
Job time : 71 secs

GenCore version 5.1.6
: Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 19:28:05 ; Search time 44 Seconds
(without alignments)
727.828 Million cell updates/sec

Title: US-10-677-179-8

Perfect score: 2200

Sequence: 1 MTTVPKPOQVAHSRARLAPR.....VAQCEPALTALQEPVGSF 429

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2200	100.0	429	US-09-743-207-8	Sequence 8, Appli
2	1797.5	81.7	440	US-09-743-207-2	Sequence 2, Appli
3	1541	70.0	431	US-09-103-331-39	Sequence 39, Appl
4	1541	70.0	431	US-09-631-594-48	Sequence 48, Appl
5	1257	57.1	307	US-09-743-207-12	Sequence 12, Appl
6	1191.5	54.2	362	US-09-103-331-38	Sequence 38, Appl
7	1191.5	54.2	362	US-09-631-594-47	Sequence 47, Appl
8	1066	48.5	257	US-09-743-207-6	Sequence 6, Appli
9	938	42.6	376	US-09-103-331-42	Sequence 42, Appl
10	938	42.6	376	US-09-631-594-51	Sequence 51, Appl
11	930	42.3	432	US-09-103-331-40	Sequence 40, Appl
12	930	42.3	432	US-09-631-594-49	Sequence 49, Appl
13	859	39.0	377	US-09-489-039A-12546	Sequence 12546, A
14	840.5	38.6	362	US-09-543-681A-4292	Sequence 4292, Ap
15	840.5	38.2	357	US-09-103-331-41	Sequence 41, Appl
16	840.5	38.2	357	US-09-631-594-50	Sequence 50, Appl
17	820.5	37.3	359	US-09-248-796A-18202	Sequence 18202, A
18	786	35.7	379	US-09-252-991A-30159	Sequence 30159, A
19	741	33.7	536	US-09-328-352-7662	Sequence 7662, Ap
20	731	33.2	536	US-09-103-331-2	Sequence 2, Appli
21	729	33.1	374	US-09-540-236-2230	Sequence 2230, Ap
22	727	33.0	536	US-09-631-594-46	Sequence 46, Appl
23	644	29.3	527	US-09-103-331-4	Sequence 4, Appli
24	644	29.3	527	US-09-631-594-55	Sequence 55, Appl
25	642.5	29.2	388	US-09-438-185A-1038	Sequence 1038, Ap
26	641	29.1	376	US-09-198-452A-1112	Sequence 1112, Ap
27	611	27.8	339	US-09-902-540-16343	Sequence 16343, A

28	478.5	21.8	168	4	US-09-743-207-14	Sequence 14, Appl
29	459.5	20.9	392	4	US-09-107-532A-4840	Sequence 4840, Ap
30	437	19.9	425	3	US-09-134-001C-3143	Sequence 3143, Ap
31	419.5	19.1	391	4	US-09-107-433-4630	Sequence 4630, Ap
32	416.5	18.9	388	4	US-09-583-110-3543	Sequence 3543, Ap
33	413.5	18.8	395	4	US-09-134-000C-4982	Sequence 4982, Ap
34	380.5	17.3	353	4	US-09-710-279-2040	Sequence 2040, Ap
35	353	16.0	108	4	US-09-743-207-4	Sequence 4, Appli
36	295	13.4	81	4	US-09-743-207-10	Sequence 10, Appl
37	102.5	4.7	432	4	US-09-252-991A-22020	Sequence 22020, A
38	101	4.6	96	4	US-09-248-796A-18203	Sequence 18203, A
39	97.5	4.4	434	4	US-09-252-991A-27401	Sequence 27401, A
40	97	4.4	233	4	US-09-328-352-5981	Sequence 5981, Ap
41	96.5	4.4	402	4	US-09-252-991A-27689	Sequence 27689, A
42	96.5	4.4	711	4	US-09-252-991A-22259	Sequence 22259, A
43	96	4.4	304	4	US-09-489-039A-13414	Sequence 13414, A
44	96	4.4	407	4	US-09-252-991A-21346	Sequence 21346, A
45	96	4.4	409	4	US-09-252-991A-21338	Sequence 21338, A

ALIGNMENTS

RESULT 1

US-09-743-207-8
; Sequence 8, Application US/09743207
; Patent No. 665331
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Pember, Stephen O.
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-A
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: PCT/US99/16353
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Zea mays
US-09-743-207-8

Query Match	Score	DB 4;	Length	429;
Best Local Similarity	100.0%;	Pred. No. 1.8e-234;		
Matches	429;	Conservative	0;	Mismatches
			0;	Indels
			0;	Gaps
				0;
Qy	1	MTTVPKPOQVAHSRARLAPRAIGALLFPAPASSLSRFAVHRCRTARLEVKASGNTFGNYF	60	
Db	1	MTTVPKPOQVAHSRARLAPRAIGALLFPAPASSLSRFAVHRCRTARLEVKASGNTFGNYF	60	
Qy	61	QVATYGSHGCGVIGCCPPRIPLTEADIQVDELRRPGOSRTITSRKETDTCILSG	120	
Db	61	QVATYGSHGCGVIGCCPPRIPLTEADIQVDELRRPGOSRTITSRKETDTCILSG	120	
Qy	121	THEGVTTGTPILVTPNTDQIGSDHREIANVYRSHADATYDFKYGVRAVQGGSSGRK	180	
Db	121	THEGVTTGTPILVTPNTDQIGSDHREIANVYRSHADATYDFKYGVRAVQGGSSGRK	180	
Qy	181	TGVRVAAGALPKKILKLCGLEILSFVSKVHQVLPEDAVDYGVTLEQIESNIVRCDDP	240	
Db	181	TGVRVAAGALPKKILKLCGLEILSFVSKVHQVLPEDAVDYGVTLEQIESNIVRCDDP	240	
Qy	241	EYAEKMDAIDRVVRGDSVCGVITCVARNVPRGIGSPVFKLSELAKMLSPASNGF	300	
Db	241	EYAEKMDAIDRVVRGDSVCGVITCVARNVPRGIGSPVFKLSELAKMLSPASNGF	300	
Qy	301	EIGSGFACTDLTSGEHNDDEFYMDKAGSVRTRTNSGGVGGISNVEIVHFKVAPKPTPSI	360	

```
Db 301 EIGSGFAGTDLTSGSHNDFYMDKAGSVKTRTRNRSQVGGISNVEIVHFKVAFKPTPSI 360
Qy 361 GVKQNTVSRRQNVELLARGHRDPCVAPRAVPVVSMAALVLMQDMAHVAQCEMFALNT 420
Db 361 GVKQNTVSRRQNVELLARGHRDPCVAPRAVPVVSMAALVLMQDMAHVAQCEMFALNT 420
Qy 421 ALQEPVGSF 429
Db 421 ALQEPVGSF 429

RESULT 2
US-09-743-207-2
; Sequence 2, Application US/09743207
; Patent No. 6653531
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Pember, Stephen O.
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-A
; CURRENT APPLICATION NUMBER: US/09/743,207
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: PCT/US99/16353
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Zea mays
US-09-743-207-2

Query Match 81.7%; Score 1797.5; DB 4; Length 440;
Best Local Similarity 81.9%; Pred. No. 6.8e-190;
Matches 349; Conservative 32; Mismatches 42; Indels 3; Gaps 2;

Qy 4 VPKPOQVAHSRARLAPRAIGALLFAPASSLSRFAV-HRCRTARLEVKASGNTFGNYFQV 62
Db 5 VSQPPVSAARSATRFPLRGIGALPESAP--TSRLSVGRRRRASSLLEVKASGNVFGNYFQV 62
Qy 63 ATYGESHGGVGCVISGCGPPRIPLTEADLQVELDRRRPGQSRTSTRKETDTCKILSGTH 122
Db 63 ATYGESHGGVGCVISGCGPPRIPLTEADMQVELDRRRPGQSRTSTRKETDTCKILSGTH 122
Qy 123 EGVTTGTPILVIPNTDQIGSDHREBIANVYRPSHADATYDFKYGVRAVQGGRRSSGRKT 182
Db 123 DGMTTGTPIHVFVNTDQGGDYSEMSKAYRPSHADATYDFKYGVRAVQGGRRSSARETI 182
Qy 183 GRVAAGALPKILKULKCGLEILSFVSKVHQVVLPEDAVDYGSVTLQETESNIVRCPPDEY 242
Db 183 GRVAAGALAKTILKLSGVEILAFYSKVHQVVLPEDAVDYETVTLTHTESNIVRCPPDEY 242
Qy 243 AEKMDAIDRVVRGDSVGGVITCVARNVPRGLSGSPVDKLESELAKAWLSIPASNGPEI 302
Db 243 AEKMTAAIDTVRVRGDSIGGVVITCARNVPRGLSGSPVDKLESELAKAWLSIPASKGPEI 302
Qy 303 GSGFAGTDLTSGSHNDFYMDKAGSVKTRTRNRSQVGGISNVEIVHFKVAFKPTPSIGV 362
Db 303 GSGFAGTDLTSGSHNDFYMDKAGSVKTRTRNRSQVGGISNVEIVHFKVAFKPTPSIGV 362
Qy 363 KQNTVSRRQNVELLARGHRDPCVAPRAVPVVSMAALVLMQDMAHVAQCEMFALNTAL 422
Db 363 KQNTVSREHEDVELLARGHRDPCVAPRAVPVVSMAALVLMQDMAHVAQCEMFALNLAL 422
Qy 423 QEPVGS 428
Db 423 QEPVGS 428

RESULT 3
```

```
US-09-103-331-39
; Sequence 39, Application US/09103331
; Patent No. 6699654
; GENERAL INFORMATION:
; APPLICANT: McLeod, Rima W.
; APPLICANT: Roberts, Craig W.
; APPLICANT: Roberts, Fiona
; APPLICANT: Johnson, Jennifer J.
; APPLICANT: Mets, Laurens J.
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS, DIAGNOSTIC
; TITLE OF INVENTION: REAGENTS, AND VACCINES BASED ON UNIQUE APICOMPLEXAN
; TITLE OF INVENTION: PARASITE COMPONENTS
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/103,331
; APPLICATION NUMBER: US/09/103,331
; FILING DATE: 23-JUN-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 8336/9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 431 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-103-331-39

Query Match 70.0%; Score 1541; DB 4; Length 431;
Best Local Similarity 72.7%; Pred. No. 1.6e-161;
Matches 287; Conservative 55; Mismatches 53; Indels 0; Gaps 0;

Qy 32 SSSLRFAVHRCRTARLEVKASGNTFGNYFOVATYGESHGCGVGVISGCPRIPLTEADL 91
Db 29 SSNLRFTHRSQPKRLBIAAGNTFGNYFRVTTFGESHGGVGCIIIDGCPPLPLSSDM 88
Qy 92 QVELDRRRPGQSRTSTRKETDTCKILSGTHGVTGTPILVIVPNTDQIGSDHREIANV 151
Db 89 QVELDRRRPGQSRTSTRKETDTCKILSGTAGDTGTPSKVKEVPNTDQGNDSLSLA 148
Qy 152 YRPSHADATYDFKYGVRAVQGGRRSGRKTGVRVAAGALPKILKLCGLEILSFVSKVH 211
Db 149 YRPSHADATYDFKYGVRSVQGGRRSSARETIGRVAAGAVAKILKLYSGTEILAYVSQVH 208
Qy 212 QVVLPEDAVDYGSVTLQETESNIVRCPPDEYAEKMDAIDRVVRGDSVGGVITCVARNV 271
Db 209 NVVLPEDLVNQIVTLQETESNIVRCNPENPEYAEKMGIGAIIDYRVRGDSVGGVITCVARNV 268
Qy 272 PRGLGSVPFVKLESELAKAWLSIPASNGFEGISGFGACTDLTSGSHNDFYMDKAGSVKTR 331
Db 269 PRGLGTFVFDKLEAEALAKAWLSIPATKGFEGSGFACTFMTGSEHNDFYMDKAGSVKTR 328
Qy 332 TNRSQVGGISNVEIVHFKVAFKPTPSIGVKQNTVSRRQNVELLARGHRDPCVAPRAV 391
Db 329 TNRSQVGGISNVEIVHFKVAFKPTPSIGVKQNTVSRRQNVELLARGHRDPCVAPRAV 388
```

QY 392 PVVESAALVLDQMAHVAQCEMFALNTALQEPV 426
Db 389 PMVEAMVALVLDQMTQVACQMLFPVNLTLQEP 423

RESULT 4

US-09-631-594-48
Sequence 48, Application US/09631594
Patent No. 6737237

GENERAL INFORMATION:

APPLICANT: MCLEOD, RIMA W.
APPLICANT: ROBERTS, CRAIG W.
APPLICANT: ROBERTS, FIONA
APPLICANT: JOHNSON, JENNIFER J.
APPLICANT: KIRLSITS, MICHAEL
APPLICANT: FERGUSON, DAVID
APPLICANT: LYONS, RUSSELL
APPLICANT: MUI, ERNEST
APPLICANT: MACK, DOUG
APPLICANT: SAMUEL, BENJAMIN
APPLICANT: GORNICKI, PIOTR
APPLICANT: ZUTHER, ELLEN

TITLE OF INVENTION: ANTIMICROBIAL AGENTS, DIAGNOSTIC REAGENTS, AND VACCINES
FILE OF INVENTION: BASED ON UNIQUE APICOMPLEXAN PARASITE COMPONENTS

FILE REFERENCE: 19338-90966

CURRENT APPLICATION NUMBER: US/09/631,594

CURRENT FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: PCT/US00/11478

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: PCT/US97/12497

PRIOR FILING DATE: 1997-07-18

NUMBER OF SEQ ID NOS: 83

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 48

LENGTH: 431

TYPE: PRT

ORGANISM: Solanum lycopersicum

US-09-631-594-48

Query Match 70.0%; Score 1541; DB 4; Length 431;

Best Local Similarity 72.7%; Pred. No. 1.6e-161;

Matches 287; Conservative 55; Mismatches 53; Indels 0; Gaps 0;

QY 32 SSSLFVHRCRTARLEKVCASGNTGNYFOVATYGESGGVGVCGVSGCPPIPLTEADL 91
Db 29 SSNLFPPTHRSQPKLEIQAGNTGNYFRTVTFGESGGVGVCGCIIDGCPPLPLSESDM 88

QY 92 QVELDRRPGQSRITSTRKETDTCKILSGTHGVTGTPIVIVPNTDQIGSDHREIANV 151
Db 89 QVELDRRPGQSRITTPKRTDTCKILSGTAGLTGSPKIVFVNTDQNGDYSEMSLA 148

QY 152 YRPSHADATYDFKYGVRAVQGGRRSGRTKTVGRVAAGALPKKILKLCGLEILSFVSKVH 211
Db 149 YRPSHADATYDFKYGVRSVQGGRRSSARETIGRAVAGAVAKKILKLYSGTEILAVVSQVH 208

QY 212 QVLPEDAVDYSVTLEQIESNIVRCPPPEYAEKMDAIDRVVRGDSVGGVITCVARNV 271
Db 209 NVLPEDLVNDQIVTLEQIESNIVRCPPPEYAEKMDAIDRVVRGDSVGGVITCVARNV 268

QY 272 PRGLSPVDFKLESELAKAMLSIPASNGFEIGSGFAGTDLTGSEHNDFFYMDKAGSVRTR 331
Db 269 PRGLTVPDFKLESELAKAMLSIPATKGFEGSGFAGTDMTGEHNDFFYMDHDIQRTK 328

QY 332 TNRSVGQVGGISNVEIVHFKVAFKPTTSGVQNTVSRERQNVLLARGHDPCVAPRAV 391
Db 329 TNRSVGQVGGISNVEIVHFKVAFKPTTSGVQNTVSRERQNVLLARGHDPCVAPRAV 388

QY 392 PVVESAALVLDQMAHVAQCEMFALNTALQEPV 426
Db 389 PMVEAMVALVLDQMTQVACQMLFPVNLTLQEP 423

RESULT 5

US-09-743-207-12
Sequence 12, Application US/09743207
Patent No. 6653531

GENERAL INFORMATION:

APPLICANT: Cahoon, Rebecca E.
APPLICANT: Falco, Saverio Carl
APPLICANT: Pember, Stephen O.

TITLE OF INVENTION: Chorismate Biosynthesis Enzymes

FILE REFERENCE: BB-1159-A

CURRENT APPLICATION NUMBER: US/09/743,207

CURRENT FILING DATE: 2001-01-04

PRIOR APPLICATION NUMBER: PCT/US99/16353

PRIOR FILING DATE: 1999-07-20

PRIOR APPLICATION NUMBER: 60/093,611

PRIOR FILING DATE: 1998-07-21

NUMBER OF SEQ ID NOS: 14

SOFTWARE: Microsoft Office 97

SEQ ID NO 12

LENGTH: 307

TYPE: PRT

ORGANISM: Oryza sativa

US-09-743-207-12

Query Match 57.1%; Score 1257; DB 4; Length 307;

Best Local Similarity 81.9%; Pred. No. 2.6e-130;

Matches 245; Conservative 22; Mismatches 26; Indels 6; Gaps 2;

QY 15 ARLAPRAIG----ALLEFAPASSLSIRFAVHRCRTARLEKVCASGNTGNYFOVATYGESHG 70
Db 11 ARVLPRGGGGGFRAPPESAPA--SLRFSVGRRAARLEVKASANVFGNYFOVATYGESHG 68

QY 71 GGVCVLSGCCPPIPLTEADLQVELDRRPGQSRITSTRKETDTCKILSGTHGVTGTCTP 130
Db 69 GGVCVLSGCCPPIPLTEADLQVELDRRPGQSRITSTRKETDTCKILSGTHGVTGTCTP 128

QY 131 ILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGVRAVQGGRRSGRTKTVGRVAAGAL 190
Db 129 IHVFPNTDQGGDYSEMAKAYRPSHADATYDFKYGVRAVQGGRRSSARETIGRVAAGAL 188

QY 191 PKKILKLCGLEILSFVSKVHVLPEDAVDYSVTLEQIESNIVRCPPPEYAEKMDAI 250
Db 189 AKKILKLSGVEILAFVSKVHVLPEDAVDYSVTLEQIESNIVRCPPPEYAEKMDAI 248

QY 251 DRVVRGDSVGGVITCVARNVPRGSGSPVDFKLESELAKAMLSIPASNGFEIGSGFAC 309
Db 249 DKVRVGDGSGVVTCTARNVPRGSGSPVDFKLESELAKAMLSIPASNGFEIGSGFVFT 307

RESULT 6

US-09-103-331-38
Sequence 38, Application US/09103331
Patent No. 6699654

GENERAL INFORMATION:

APPLICANT: McLeod, Rima W.
APPLICANT: Roberts, Craig W.
APPLICANT: Roberts, Fiona
APPLICANT: Johnson, Jennifer J.

APPLICANT: Mets, Laurens J.

TITLE OF INVENTION: ANTIMICROBIAL AGENTS, DIAGNOSTIC

TITLE OF INVENTION: REAGENTS, AND VACCINES BASED ON UNIQUE APICOMPLEXAN

TITLE OF INVENTION: PARASITE COMPONENTS

NUMBER OF SEQUENCES: 44

CORRESPONDENCE ADDRESS:

ADDRESSEE: BRINKS, HOFER, GILSON & LIONE

STREET: NBC Tower - Suite 3600, 455 N. Cityfront

STREET: Plaza Drive

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60611-5599

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

```
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/103.331
; FILING DATE: 23-JUN-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 8336/9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-103-331-38

Query Match          54.2%; Score 1191.5; DB 4; Length 362;
Best Local Similarity 62.2%; Pred. No. 6.1e-123;
Matches 225; Conservative 57; Mismatches 77; Indels 3; Gaps 1;

QY 53 GNTFGNYQVATYGESHGCGVCGVISCPPRIPLTEADLOVELDRRRPGQSRTSTRKET 112
DB 2 GNTFGSLFRITTFGESHGCGVVIIDGCPRLAISPEEIQVLDLDRRRPGQSRTSTRKET 61
QY 113 DTCKILSGTHEGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGVRAVQ 172
DB 62 DQCEILSGVFEKTLGTPIAILVRNKDARSQDYNEMAVKYRPSHADATYEAKEYGIRNWQ 121
QY 173 GGRSSGRKTVGRVAAGALPKKILKLCGLEILSFVSKVHQVVLPEDAVDYGSVTLEQIES 232
DB 122 GGRSSARETIGRVAAGAIKLLAQFNGVEIVAVYKSIQDI---EATVDSNTVTLEQVES 178
QY 233 NIVRCPDPEYAEKMDAIDRVVRVGDVSGVITCVARNVPRGLSGSPVDPKLESELAKAML 292
DB 179 NIVRCPDECAEKMIERIDQVLRQKDSIGGVVECAIRNAPKGLGEPVDFDKLEADLAKAM 238
QY 293 SIPASNGEIGSGFAGTDLTGSEHNDEFYMDKAGSVTRTRNSGGVQGGISNVEIVHFKV 352
DB 239 SLPATKGFEFGSGFAGTLLTGSOHNDEYILDEAGEWRTTRNSGGVQGGISNGEPIIMRI 298
QY 353 AFKPTPSIGVKQNTVSRBRQNVVELLARGHDPVAPRAVPVVSMAALVLMQDLMAHVAQ 412
DB 299 AFKPTATIGQEQKTVSNIGETTLAAGKGRHDPVLPRAVPVVEAMAALVLCDDLRRFOAQ 358
QY 413 CE 414
DB 359 CK 360

RESULT 7
US-09-631-594-47
; Sequence 47, Application US/09631594
; Patent No. 6737237
; GENERAL INFORMATION:
; APPLICANT: MCLEOD, RIMA W.
; APPLICANT: ROBERTS, CRAIG W.
; APPLICANT: ROBERTS, FIONA
; APPLICANT: JOHNSON, JENNIFER J.
; APPLICANT: KIRISITS, MICHAEL
; APPLICANT: FERGUSON, DAVID
; APPLICANT: LYONS, RUSSELL
; APPLICANT: MUI, ERNEST
; APPLICANT: MACK, DOUG
; APPLICANT: SAMUEL, BENJAMIN
; APPLICANT: GORNICKI, PIOTR
; APPLICANT: ZUTHER, ELLEN

;
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS, DIAGNOSTIC REAGENTS, AND VACCINES
; TITLE OF INVENTION: BASED ON UNIQUE APICOMPLEXAN PARASITE COMPONENTS
; FILE REFERENCE: 19338-90966
; CURRENT APPLICATION NUMBER: US/09/631,594
; CURRENT FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: PCT/US00/11478
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: PCT/US97/12497
; PRIOR FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Synecchocystis sp.
; US-09-631-594-47

Query Match          54.2%; Score 1191.5; DB 4; Length 362;
Best Local Similarity 62.2%; Pred. No. 6.1e-123;
Matches 225; Conservative 57; Mismatches 77; Indels 3; Gaps 1;

QY 53 GNTFGNYQVATYGESHGCGVCGVISCPPRIPLTEADLOVELDRRRPGQSRTSTRKET 112
DB 2 GNTFGSLFRITTFGESHGCGVVIIDGCPRLAISPEEIQVLDLDRRRPGQSRTSTRKET 61
QY 113 DTCKILSGTHEGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGVRAVQ 172
DB 62 DQCEILSGVFEKTLGTPIAILVRNKDARSQDYNEMAVKYRPSHADATYEAKEYGIRNWQ 121
QY 173 GGRSSGRKTVGRVAAGALPKKILKLCGLEILSFVSKVHQVVLPEDAVDYGSVTLEQIES 232
DB 122 GGRSSARETIGRVAAGAIKLLAQFNGVEIVAVYKSIQDI---EATVDSNTVTLEQVES 178
QY 233 NIVRCPDPEYAEKMDAIDRVVRVGDVSGVITCVARNVPRGLSGSPVDPKLESELAKAML 292
DB 179 NIVRCPDECAEKMIERIDQVLRQKDSIGGVVECAIRNAPKGLGEPVDFDKLEADLAKAM 238
QY 293 SIPASNGEIGSGFAGTDLTGSEHNDEFYMDKAGSVTRTRNSGGVQGGISNVEIVHFKV 352
DB 239 SLPATKGFEFGSGFAGTLLTGSOHNDEYILDEAGEWRTTRNSGGVQGGISNGEPIIMRI 298
QY 353 AFKPTPSIGVKQNTVSRBRQNVVELLARGHDPVAPRAVPVVSMAALVLMQDLMAHVAQ 412
DB 299 AFKPTATIGQEQKTVSNIGETTLAAGKGRHDPVLPRAVPVVEAMAALVLCDDLRRFOAQ 358
QY 413 CE 414
DB 359 CK 360

RESULT 8
US-09-743-207-6
; Sequence 6, Application US/09743207
; Patent No. 6653531
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Pember, Stephen O.
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-A
; CURRENT APPLICATION NUMBER: US/09/743,207
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: PCT/US99/16353
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Triticum aestivum
; US-09-743-207-6
```

Query Match 48.5%; Score 1066; DB 4; Length 257;
Best Local Similarity 82.7%; Pred. No. 2.7e-109;
Matches 201; Conservative 27; Mismatches 15; Indels 0; Gaps 0;
QY 186 AAGALPKKILKCGLEILSFVSKVHQVLPEDAVDYGVTLEQIESNIVRCPDPEYAK 245
DB 3 AAGAVAKTILKCGVEILAFVSKVHQVLPEDAVDYETLTLDQIESNIVRCPDPEYAK 62
QY 246 MIDAIDRVVRGDSVGGVITCVARNVPRGLGSPVDFKLESELAAMLSIPASNGEIGSG 305
DB 63 MIDAIDKVRNGSIGGVVTCIARNVPRGLGSPVDFKLESELAAMLSIPASNGEIGSG 122
QY 306 FAGTDLTSGSEHDEFTYMDKAGSVRTNRSGVGGVQGGISNVEIVHFKVAFKPTPSIGVKON 365
DB 123 FAGTDLTSGSEHDEFTYMDKAGSVRTNRSGVGGVQGGISNVEIVHFKVAFKPTPSIGVKON 182
QY 366 TVSRERQNVELLARGHDCVAPRAVPVVSMAALVMDQLMAHVAQCMBFALNTALQEP 425
DB 183 TVTROHEDIETLLTRGRHDCVAPRAVPVVSMAALVMDQLMAHVAQCMBFALNTALQEP 242
QY 426 VGS 428
DB 243 IGS 245

RESULT 9
US-09-103-331-42
; Sequence 42, Application US/09103331
; Patent No. 6699654
; GENERAL INFORMATION:
; APPLICANT: McLeod, Rima W.
; APPLICANT: Roberts, Craig W.
; APPLICANT: Roberts, Fiona
; APPLICANT: Johnson, Jennifer J.
; APPLICANT: Metz, Laurens J.
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS, DIAGNOSTIC
; TITLE OF INVENTION: REAGENTS, AND VACCINES BASED ON UNIQUE APICOMPLEXAN
; TITLE OF INVENTION: PARASITE COMPONENTS
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOPFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/103,331
; FILING DATE: 23-JUN-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 8336/9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-103-331-42

Query Match 42.6%; Score 938; DB 4; Length 376;
Best Local Similarity 49.9%; Pred. No. 7.6e-95;
Matches 183; Conservative 67; Mismatches 109; Indels 8; Gaps 3;
QY 54 NTFNGYFOVATYGGSHGGVGVSGCPPIPLTEADLQVELDRRRPQOSRITSTRKETD 113
DB 2 STFGKLFVTTYGGSHCKSVGCIVDGVPFGMSLTEADIQPOLTRRRPQOSKLSRPRDEKD 61
QY 114 TCKILSGTHGVTGTPILVVPNTDOIGSDHREIFANVYRPSHADATYDFKYGVRAVQGG 173
DB 62 RVBQSGTFEGTKTLGTPIAMIKNEDQRPDYSMDKFPFSPSHADFTYSEKYGKIKASGG 121
QY 174 GRSSGRKTVGRVAAGALPKKILKCGLEILSFVSKVHQVLPEDAVD-----YGSVTL 227
DB 122 GRASARETIGRVASGAIKFLAQNSNVEIVAFVTVQICEIKMNRDSFDPBFQHLNLTITR 181
QY 228 EQLSES-NIVRCPDPEYAKMIDAI DRVVRGDSVGGVITCVARNVPRGLGSPVDFKLESE 286
DB 182 EKVDSMGPIRCPDASVAGLMVKEIKYRGKDSIGGVVTCVVRNLPTGLGEPDFKLEAM 241
QY 287 LAKAMLSTPASNGEIGSGFAGTDLTSGSEHDEFTYMDK-AGSVRTRTNRSVGGVQGGISNV 345
DB 242 LAHAMLSTPASNGEIGSGFAGTDLTSGSEHDEFTYMDK-AGSVRTRTNRSVGGVQGGISNV 301
QY 346 EIVHFKVAFKPTPSIGVKONTVSRRERQNVELLARGHDCVAPRAVPVVSMAALVMDQ 405
DB 302 ENIVSFVPFKSVATISQEQKTATYDGBEGILAAKGRHDPATVTPRAIPVEAMTALVLADA 361
QY 406 LMAHVAQ 412
DB 362 LLIQKAR 368

RESULT 10
US-09-631-594-51
; Sequence 51, Application US/09631594
; Patent No. 6737237
; GENERAL INFORMATION:
; APPLICANT: McLeod, Rima W.
; APPLICANT: Roberts, Craig W.
; APPLICANT: Roberts, Fiona
; APPLICANT: Johnson, Jennifer J.
; APPLICANT: Kirisits, Michael
; APPLICANT: Ferguson, David
; APPLICANT: Lyons, Russell
; APPLICANT: Mui, Ernest
; APPLICANT: Mack, Doug
; APPLICANT: Samuel, Benjamin
; APPLICANT: Gornicki, Piotr
; APPLICANT: Zuther, Ellen
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS, DIAGNOSTIC REAGENTS, AND VACCINES
; TITLE OF INVENTION: BASED ON UNIQUE APICOMPLEXAN PARASITE COMPONENTS
; FILE REFERENCE: 19338-90966
; CURRENT APPLICATION NUMBER: US/09/631,594
; CURRENT FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: PCT/US00/11478
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: PCT/US97/12497
; PRIOR FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 51
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-631-594-51

Query Match 42.6%; Score 938; DB 4; Length 376;
Best Local Similarity 49.9%; Pred. No. 7.6e-95;
Matches 183; Conservative 67; Mismatches 109; Indels 8; Gaps 3;
QY 54 NTFNGYFOVATYGGSHGGVGVSGCPPIPLTEADLQVELDRRRPQOSRITSTRKETD 113
DB 2 STFGKLFVTTYGGSHCKSVGCIVDGVPFGMSLTEADIQPOLTRRRPQOSKLSRPRDEKD 61

Db 2 STFGKLFRTVTTGESHCKSVGCIVDGVPFGMSLTEADIQPOLTRRRPGQSKLSTPRDEKD 61
Qy 114 TKILSGTHGVTGTPILVIVPNTDQIGSDHREITANVVRSHADATYDFKYGVRVAVOGG 173
Db 62 RVEIQSGTEFGKLTGTPIAMMIKNDQRPDHYSDMDKFPFSPHADFTYSEKYGKASSGG 121
Qy 174 GRSSGRKTVGRVAAGALPKKILKLCGLEILSFVSKVHVVLPEDAVD-----YGSVTL 227
Db 122 GRASARETIGRVAAGAIKFLAQNSEIVAFVTOIGEIKNRDSFDPFQHLNITIR 181
Qy 228 EQIES-NIVRCPDPEYAEKMDAIDRVRVRGDSVGGVITCVARNVPRGLSGVDFDKLESE 286
Db 182 EKVDWGPICPDASVAGLWKEIEKRYGNKDSIGSVVTCVVRNLTGSGPFCFKLEAM 241
Qy 287 LAKAMLSIPASNGFPIGSGFAGTDLTGSEHNDEFMDK-AGSVRTRTRNSGVQGGISNV 345
Db 242 LAHAMLSIPASKFPIGSGFQVSGVPSGKNDPPYFEKETNRLRTKTNNSGVQGGISNG 301
Qy 346 EIVHFKVAFKPTPSIGVKQNTVSRERONVELLARGHDPVAPRAVPVSVESMAALVLMQ 405
Db 302 ENIYFSVPFKSVATISQSKTATYDGBEGILAAKGRHDPVATPRAIPVEAMTALVLADA 361
Qy 406 LMAHVAQ 412
Db 362 LLIQKAR 368

RESULT 11
US-09-103-331-40
; Sequence 40, Application US/09103331
; Patent No. 6699654
; GENERAL INFORMATION:
; APPLICANT: McLeod, Rima W.
; APPLICANT: Roberts, Craig W.
; APPLICANT: Roberts, Fiona
; APPLICANT: Johnson, Jennifer J.
; APPLICANT: Metz, Laurens J.
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS, DIAGNOSTIC
; TITLE OF INVENTION: REAGENTS, AND VACCINES BASED ON UNIQUE APICOMPLEXAN
; TITLE OF INVENTION: PARASITE COMPONENTS
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; STREET: Plaza Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/103,331
; FILING DATE: 23-JUN-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 8336/9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-103-331-40

Query Match 42.3%; Score 930; DB 4; Length 432;
Best Local Similarity 48.9%; Pred. No. 7.3e-94;
Matches 201; Conservative 64; Mismatches 110; Indels 36; Gaps 7;
Qy 54 NTFGNVQVATYGBSHGGVCGVISCPCPRIPRIPLTEADLQVELDRRRPGQSRITSTRKETD 113
Db 2 STFGHYFRVTTYGESHCKSVGCIVDGVPFGMELTEDDIOQMTRRRPGQSAITTPRDEKD 61
Qy 114 TKILSGTHGVTGTPILVIVPNTDQIGSDH-REIANVY-RPSHADATYDFKYGVRVAVQ 171
Db 62 RVIIFSQTERTGVTGTPIGMLVMNEDQPKDYGNKTMDIYRPSHADWTYLEKYGVRKASS 121
Qy 172 GGGSSGRKTVGRVAAGALPKKILKLCGLEILSFVSKVHVVL-----PED 218
Db 122 GGGSSARETIGRVAAGAIKFLAQNSEIVAFVTOIGEIKNRDSFDPFQHLNITIR 181
Qy 219 ADVYGSVTLQIIESNI-VRCPDPEYAEKMDAIDRVRVRGDSVGGVITCVARNVPRGLGS 277
Db 182 LKLVNSITRETVDVSFLPVRCPDAEANKMEDLITKFPDNHDSIGGTVTCVIRNVPSGLGE 241
Qy 278 PVFDKLESELAAMLSIPASNGFPIGSGFAGTDLTGSEHNDEFY-----MDKAGS 327
Db 242 PAFDKLEAMLAHAMLSIPATKGFVSGFGGCEVPSGIHNDPPFVSAENTEIPPSVAASGA 301
Qy 328 VR-----TRNRSRGVQGGISNVIEIVHFKVAFKPTPSIGVKQNTVSRERONVELL- 377
Db 302 ARNGIPRPKLTUTTNFGSGIQQGINSAGPIYRFRVGFKAATIGOEQTATYDGTSGVLA 361
Qy 378 ARGHRDPCVAPRAVPVSVESMAALVLMQDMAHVAQCEMFAINTALQBPVGS 428
Db 362 AKGRHDSVVPRAVPIVEAMALVIMDAVLAHEARVTAKSLPLPKQTINS 412

RESULT 12
US-09-631-594-49
; Sequence 49, Application US/09631594
; Patent No. 6737237
; GENERAL INFORMATION:
; APPLICANT: McLEOD, RIMA W.
; APPLICANT: ROBERTS, CRAIG W.
; APPLICANT: ROBERTS, FIONA
; APPLICANT: JOHNSON, JENNIFER J.
; APPLICANT: KRISITS, MICHAEL
; APPLICANT: FERGUSON, DAVID
; APPLICANT: LYONS, RUSSELL
; APPLICANT: MUL, ERNEST
; APPLICANT: MACK, DOUG
; APPLICANT: SAMUEL, BENJAMIN
; APPLICANT: GORNICKI, PIOTR
; APPLICANT: ZUTHER, ELLEN
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS, DIAGNOSTIC REAGENTS, AND VACCINES
; TITLE OF INVENTION: BASED ON UNIQUE APICOMPLEXAN PARASITE COMPONENTS
; FILE REFERENCE: 19338-90966
; CURRENT APPLICATION NUMBER: US/09/631,594
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: PCT/US00/11478
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: PCT/US97/12497
; PRIOR FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 49
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Neurospora crassa
US-09-631-594-49

Query Match 42.3%; Score 930; DB 4; Length 432;
Best Local Similarity 48.9%; Pred. No. 7.3e-94;
Matches 201; Conservative 64; Mismatches 110; Indels 36; Gaps 7;
Qy 54 NTFGNVQVATYGBSHGGVCGVISCPCPRIPRIPLTEADLQVELDRRRPGQSRITSTRKETD 113

Db 2 STFGHYFRVTVYGESCHKSGVCIQVDPVPGMELTDDIQWTRRRPGQSAITTPREXQ 61
QY 114 TCKILSGTHEGVTGTPILVIVPNTDOIGSDH-REIANYV-RPSHADATYDKYGVRAVQ 171
Db 62 RVTIQSGTEGVTGTPIGMLVNNEDQPPKQYGNKTMIDYPRPSHADWTYLEKYGVKASS 121
QY 172 GGRSSGRKTVGRVAAGALPKILKLCGLLEILSFVSKVHVVL-----PRD 218
Db 122 GGRSSARETIGRVAAGATAEKYLKPRYGVVEIVAFVSSVSGSEHLPPPTAEHSPSPNPEF 181
QY 219 ADVYGSVTLEQIESNI-VRCPPDEYAEKMDAIDRVVRGDSVGGVITCVARNVPRGLGS 277
Db 182 LKLVNSITRETVDSEFLPVRCPDAEANKRMEDLITKFRDNHDSIGGTVTCVIRNVPSGLGE 241
QY 278 PVPDKLESELAKAMLSIPASNGFEIGSGFAGTDLTGSEHNDEFY-----MDKAGS 327
Db 242 PAPDKLEAMLAHAMSIPATKGFVSGSGGCEVPGSIHNDPFVSAENTEIPPSVAASGA 301
QY 328 VR-----TRNRSQVGGISNVEIVHFVKVAFKPTPSIGVKQNTVSRERQNVLL- 377
Db 302 ARNGIPRPKLTXTNFSGGIQQGINSAGPIYFRVGFKPAATIGQEGTATYDGTSEGULA 361
QY 378 AKGRHDPVAPRAVPVSMALVMDQMAHVAOCFEMPALNTALQEPVGS 428
Db 362 AKGRHDPVAPRAVPVSMALVMDVLAHEARVTAKSLLPLKQNTINS 412

RESULT 13
US-09-489-039A-12546
; Sequence 12546, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12546
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12546

Query Match 39.0%; Score 859; DB 4; Length 377;
Best Local Similarity 47.7%; Pred. No. 4.2e-86;
Matches 180; Conservative 65; Mismatches 112; Indels 20; Gaps 6;
QY 40 HRGRTRARLEKASGNTFGNYFOVATYGESHGCGVISCPCPRIPLTADQLVBLDRRR 99
Db 6 HANITKIGAVMAGNTIGOLFRTVTFGESHGLALGCIQVDPGPPGIPLTADQLVBLDRRR 65
QY 100 PGOSRTSTRKETDCKILSGTHEGVTGTPILVIVPNTDOIGSDHREIANYVRSHADA 159
Db 66 PGTSTRTTORREPDQVKILSGVFEGVTGTSIGLLIENTDQRSQDYGAIKOLFREPCHADY 125
QY 160 TYDFKYGVRVAVGGGRSSGRKTVGRVAAGALPKILKLCGLLEILSFVSKVHVVLPEDA 219
Db 126 TYEQKYGLRDYGGGRSSARETAMRVAAGAIKYLKPRYGVVEIVAFVSSVSGSEHLPPPTAEHSPSPNPEF 183
QY 220 VDYGSVTLEQIESNIVRCPPDEYAEKMDAIDR-----VVRGDSVGGVITCVARNVPRGL 275
Db 184 KDW-----DQVQNPFFCDDPD-----KIDALBELMRGLKKGDSIGAKVTVDGVPPGL 234
QY 276 GSPVFDKLESELAKAMLSIPASNGFEIGSGFAGTDLTGSEHNDEFYMDKAGSVRTRTWS 335
Db 235 GEPVFDRLDADIAHALMSINAVKVGIEGDFVFWKLRGSENDE--ITKAG---FQSNHA 289
QY 336 GGVQGGISNVEIVHFVKVAFKPTPSIGVKQNTVSRERQNVLLARGRHDPVAPRAVPVVE 395

Db 290 GGLIGSSGQIVANIALKPTSSITVPGHINRBEVEMITKGRHDPVCGIRAPVIAE 349
QY 396 SMAALVLMQDMAHVAQ 412
Db 350 AMLAIVLMDHFMQRQAQ 366
RESULT 14
US-09-543-681A-4292
; Sequence 4292, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4292
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4292

Query Match 38.6%; Score 850; DB 4; Length 362;
Best Local Similarity 46.1%; Pred. No. 3.9e-85;
Matches 167; Conservative 72; Mismatches 111; Indels 12; Gaps 3;
QY 52 SGNTFGNYFOVATYGESHGCGVISCPCPRIPLTADQLVBLDRRRPGOSRTSTRKE 111
Db 3 AGNSIGOLFRTVTFGESHGTLGCIQVDPGPPGIPLTADQLVBLDRRKPGTSRYTTORRE 62
QY 112 TDTCKILSGTHEGVTGTPILVIVPNTDOIGSDHREIANYVRPSHADATYDKYGVRAVQ 171
Db 63 PDQVRILSGVNGVTGTSIGLLIENTDQRSQDYSEIKDVFPRPGHADTYEQKYLGRDYR 122
QY 172 GGRSSGRKTVGRVAAGALPKILKLCGLLEILSFVSKVHVVLPEDAVYGSVTLEQIE 231
Db 123 GGRSSARETAMRVAAGAIKYLKPRYGVVEIVAFVSSVSGSEHLPPPTAEHSPSPNPEF 175
QY 232 SNIVRCPPDEYAEKMDAIDRVVRGDSVGGVITCVARNVPRGLSPVFDKLESELAKAM 291
Db 176 TNPPFCDDPDPSRLDALDEYMRALKKEGNSIGAKVTVAEGVPAGLGEVFDRLDADIAHAL 235
QY 292 LSIPASNGFEIGSGFAGTDLTGSEHNDEFYMDKAGSVRTRNRSQVGGVIGSINVEIVHF 351
Db 236 MSINAVKVGIEGDFVVTGTLGTEENREITKEGFS-----NHAGGVLGIGSSGQPIAH 290
QY 352 VAPKPTPSIGVKQNTVSRERQNVLLARGRHDPVAPRAVPVSMALVMDQDMAHVA 411
Db 291 IALKPTSSITTAGKTLNRQGBEVDITKGRHDPVCGIRAPVIAEAMVAIVLMDHALRQ 350
QY 412 QC 413
Db 351 QC 352

RESULT 15
US-09-103-331-41
; Sequence 41, Application US/09103331
; Patent No. 6699654
; GENERAL INFORMATION:
; APPLICANT: McLeod, Rima W.
; APPLICANT: Roberts, Craig W.
; APPLICANT: Roberts, Fiona
; APPLICANT: Johnson, Jennifer J.
; APPLICANT: Mets, Laurens J.
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS, DIAGNOSTIC
; REAGENTS, AND VACCINES BASED ON UNIQUE APICOMPLEXAN

Search completed: August 25, 2005, 19:32:30
Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 19:28:51 ; Search time 161 Seconds
(without alignments)
1046.990 Million cell updates/sec

Title: US-10-677-179-8

Perfect score: 2200

Sequence: 1 MTTVPKQQAHSRARLAPR.....VAQCEMFALNTALQEPVGSF 429

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1767149 seqs, 392926209 residues

Total number of hits satisfying chosen parameters: 1767149

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US10F_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep:*
- 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2200	100.0	429	15	US-10-677-179-8
2	2107	95.8	497	16	US-10-425-115-323000
3	2016.5	91.7	408	15	US-10-425-114-69504
4	2011.5	91.4	402	16	US-10-425-115-322998
5	1800.5	81.8	440	15	US-10-425-115-356153
6	1797.5	81.7	440	15	US-10-677-179-2
7	1790.5	81.4	452	15	US-10-425-114-57857
8	1724	78.4	426	16	US-10-437-963-201690
9	1541	70.0	431	14	US-10-268-611-39
10	1541	70.0	431	18	US-10-824-194-48
11	1527	69.4	440	14	US-10-268-611-63

12	1512	68.7	435	15	US-10-424-599-144653	Sequence 144653, A
13	1512	68.7	438	15	US-10-425-114-52276	Sequence 52276, A
14	1454	66.1	436	16	US-10-696-616-28	Sequence 28, Appl
15	1257	57.1	307	15	US-10-677-179-12	Sequence 12, Appl
16	1191.5	54.2	362	14	US-10-268-611-38	Sequence 38, Appl
17	1191.5	54.2	362	14	US-10-268-611-64	Sequence 64, Appl
18	1191.5	54.2	362	15	US-10-369-493-2677	Sequence 2677, Ap
19	1191.5	54.2	362	18	US-10-824-194-47	Sequence 47, Appl
20	1186.5	53.9	362	15	US-10-369-493-18868	Sequence 18868, A
21	1182.5	53.8	362	15	US-10-369-493-20210	Sequence 20210, A
22	1066	48.5	257	15	US-10-677-179-6	Sequence 6, Appli
23	1035	47.0	360	15	US-10-369-493-21053	Sequence 21053, A
24	981	44.6	356	15	US-10-369-493-10419	Sequence 10419, A
25	938	42.6	376	14	US-10-268-611-42	Sequence 42, Appl
26	938	42.6	376	15	US-10-369-493-21922	Sequence 21922, A
27	938	42.6	376	18	US-10-824-194-51	Sequence 51, Appl
28	934	42.5	378	17	US-10-741-849-7185	Sequence 7185, Ap
29	930	42.3	432	14	US-10-268-611-40	Sequence 40, Appl
30	930	42.3	432	18	US-10-824-194-49	Sequence 49, Appl
31	927.5	42.2	285	15	US-10-424-599-144654	Sequence 144654, A
32	903	41.0	240	16	US-10-425-115-356150	Sequence 356150, A
33	888.5	40.4	358	15	US-10-282-122A-48831	Sequence 48831, A
34	886.5	40.3	473	15	US-10-369-493-3972	Sequence 3972, Ap
35	862	39.2	361	15	US-10-282-122A-78247	Sequence 78247, A
36	860	39.1	361	9	US-09-815-242-10232	Sequence 10232, A
37	860	39.1	361	15	US-10-369-493-23483	Sequence 23483, A
38	860	39.1	361	15	US-10-282-122A-56623	Sequence 56623, A
39	858	39.0	352	15	US-10-282-122A-61027	Sequence 61027, A
40	857	39.0	361	9	US-09-815-242-11708	Sequence 11708, A
41	856	38.9	361	14	US-10-268-611-65	Sequence 65, Appl
42	855	38.9	361	15	US-10-282-122A-59797	Sequence 59797, A
43	850	38.6	361	15	US-10-282-122A-69065	Sequence 69065, A
44	846.5	38.5	357	15	US-10-282-122A-66978	Sequence 66978, A
45	843	38.3	360	15	US-10-369-493-21155	Sequence 21155, A

ALIGNMENTS

RESULT 1
US-10-677-179-8
; Sequence 8, Application US/10677179
; Publication No. US20040082050A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Pember, Stephen O.
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-A
; CURRENT APPLICATION NUMBER: US/10/677,179
; PRIOR FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: US/09/743,207
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: PCT/US99/16353
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Zea mays
US-10-677-179-8

Query Match 100.0%; Score 2200; DB 15; Length 429;
Best Local Similarity 100.0%; Pred. No. 1.1e-209;
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MTTVPKQQAHSRARLAPRAIGALLEFAFPASSLSRFAVHRCRTARLEKASGNTFGNYF 60
Db 1 MTTVPKQQAHSRARLAPRAIGALLEFAFPASSLSRFAVHRCRTARLEKASGNTFGNYF 60

```
QY 61 QVATYGESHGVCVISCPPRIPLTEADLQVELDRRRPGQSRTSTRKETDTCKILSG 120
Db 61 QVATYGESHGVCVISCPPRIPLTEADLQVELDRRRPGQSRTSTRKETDTCKILSG 120
QY 121 THEGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGVAQVGGGRSSGRK 180
Db 121 THEGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGVAQVGGGRSSGRK 180
QY 181 TVGRVAAGALPKKILKLCGLEILSFVSKVQHVLPEDAVDYGSVTLQIESNIVRCDDP 240
Db 181 TVGRVAAGALPKKILKLCGLEILSFVSKVQHVLPEDAVDYGSVTLQIESNIVRCDDP 240
QY 241 EYAEKMDAIDRVRVRGDSVGGVITCVARNVPRGLGSPVDFDKLESELAKAMLSIPASNGF 300
Db 241 EYAEKMDAIDRVRVRGDSVGGVITCVARNVPRGLGSPVDFDKLESELAKAMLSIPASNGF 300
QY 301 EIGSGFAGTDLTGSEHNDEFYMDKAGSVRTRNRRSGVQGGISNVEIVHFVKVAFKPTPSI 360
Db 301 EIGSGFAGTDLTGSEHNDEFYMDKAGSVRTRNRRSGVQGGISNVEIVHFVKVAFKPTPSI 360
QY 361 GVKQNTVSRERQNVLLARGRHDPCVAPRAVPVSVESMAALVLMQJMAHVAQCEMFALNT 420
Db 361 GVKQNTVSRERQNVLLARGRHDPCVAPRAVPVSVESMAALVLMQJMAHVAQCEMFALNT 420
QY 421 ALQEPVGSF 429
Db 421 ALQEPVGSF 429
```

RESULT 2

```
US-10-425-115-323000
; Sequence 323000, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222) B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 323000
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(497)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_57640C.1.pep
US-10-425-115-323000
```

Query Match 95.8%; Score 2107; DB 16; Length 497;

Best Local Similarity 84.9%; Pred. No. 2.5e-200;

Matches 422; Conservative 2; Mismatches 5; Indels 68; Gaps 2;

```
QY 1 MTTVPKQQAHSRRLAPRAIGALLEFAPASSLSRFAVHRCRTARLEVKASGNTFGNYF 60
Db 1 MTTVPKQQAHSRRLAPRAIGALLEFAPASSLSRFAVHRCRTARLEVKASGNTFGNYF 60
QY 61 QVATYGESHGVCVISCPPRIPLTEADLQVELDRRRPGQSRTSTRKETDTCKILSG 120
Db 61 QVATYGESHGVCVISCPPRIPLTEADLQVELDRRRPGQSRTSTRKETDTCKILSG 120
QY 121 THEGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGVAQVGGGRSSGRK 180
Db 121 THEGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGVAQVGGGRSSGRK 180
QY 181 TVGRVAAGALPKKILKLCGLEILSFVSKVQHVLPEDAVDYGSVTLQIESNIVRCDDP 240
```

```
Db 181 TVGRVAAGALPKKILKLCGLEILSFVSKVQHVLPEDAVDYGSVTLQIESNIVRCDDP 240
QY 241 EYAEKMDAIDRVRVRGDSVGGVITCVARNVPRGLGSPVDFDKLESELAKAMLSIPASNGF 300
Db 241 EYAEKMDAIDRVRVRGDSVGGVITCVARNVPRGLGSPVDFDKLESELAKAMLSIPASNGF 300
QY 301 EIGSGFAGTDLTGSEHNDEFYMDKAGSVRTRNRRSGVQGGISNVEIVHFVKVAFKPTPSI 360
Db 301 EIGSGFAGTDLTGSEHNDEFYMDKAGSVRTRNRRSGVQGGISNVEIVHFVKVAFKPTPSI 360
QY 361 NGIGLFAVXVQGGISNGISVHFVKVAFKPTPSIGVKQNTVSRERQNVLLARGRHDPCVAPR 420
QY 390 -----AVPVESMAALVLMQJMAHVAQ 412
Db 421 XHVRVMTLLAASATILFNFFPATATATAIRGLLLTWLXAVPVESMAALVLMQJMAHVAQ 480
QY 413 CEMFALNTALQEPVGSF 429
Db 481 CEMFALNTALQEPVGSF 497

RESULT 3
US-10-425-114-69504
; Sequence 69504, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313) B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 69504
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73045G02_FLI.pep
US-10-425-114-69504
```

```
Query Match 91.7%; Score 2016.5; DB 15; Length 408;
Best Local Similarity 92.8%; Pred. No. 1.9e-191;
Matches 398; Conservative 1; Mismatches 3; Indels 27; Gaps 1;

QY 1 MTTVPKQQAHSRRLAPRAIGALLEFAPASSLSRFAVHRCRTARLEVKASGNTFGNYF 60
Db 7 MTTVPKQQAHSRRLAPRAIGALLEFAPASSLSRFAVHRCRTARLEVKASGNTFGNYF 66
QY 61 QVATYGESHGVCVISCPPRIPLTEADLQVELDRRRPGQSRTSTRKETDTCKILSG 120
Db 67 QVATYCESHGVCVISCPPRIPLTEADLQVELDRRRPGQSRTSTRKETDTCKILSG 126
QY 121 THEGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGVAQVGGGRSSGRK 180
Db 127 THEGVTGTPILVIVPNTDQIGSDHREIANVYRPSHADATYDFKYGVAQVGGGRSSGRK 186
QY 181 TVGRVAAGALPKKILKLCGLEILSFVSKVQHVLPEDAVDYGSVTLQIESNIVRCDDP 240
Db 187 TVGRVAAGALPKKILKLCGLEILSFVSKVQHVLPEDAVDYGSVTLQIESNIVRCDDP 219
QY 241 EYAEKMDAIDRVRVRGDSVGGVITCVARNVPRGLGSPVDFDKLESELAKAMLSIPASNGF 300
Db 220 EYAEKMDAIDRVRVRGDSVGGVITCVARNVPRGLGSPVDFDKLESELAKAMLSIPASNGF 279
```



```
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: PCT/US99/16353
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Zea mays
US-10-677-179-2

Query Match      81.7%; Score 1797.5; DB 15; Length 440;
Best Local Similarity 81.9%; Pred. No. 1.3e-169;
Matches 349; Conservative 32; Mismatches 42; Indels 3; Gaps 2;

Qy  4 VPKPOQVAHSRARLAPRAIGALLEFAPASSLSRFVAV-HRCRTARLEVKASGNTFGNYFQV 62
Db  5 VSQPPVSARASTRFLPRGIGALPESAP--TSLRLSVGRRRAASLEVKASGNVFGNYFQV 62
Qy  63 ATYGESHGGVGCVISGCPPIPLTEADLQVELDRRRPGQSRIITSTRKETDTCKILSGTH 122
Db  63 ATYGESHGGVGCVISGCPPIPLTEADLQVELDRRRPGQSRIITSTRKETDTCKILSGTH 122
Qy  123 EGVTTGTPILVIVPNTDOIGSDHREIANVVRPSSHADATYDFKYGVRVAVQGGRRSGRTV 182
Db  123 DGMTTGTPIHVPNTDQGGDYSEMSKAYRPSHADATYDFKYGVRVAVQGGRRSSARETI 182
Qy  183 GRVAAGALPKKILKLCGLLEILSFVSKVHQVLPEDAVDYGSVTLQIESNIVRCPPDEY 242
Db  183 GRVAAGALAKKILKLSGVLEILAFVSKVHQVLPEDAVDYETVTLHIESNIVRCPPDEY 242
Qy  243 AEKMDAIDRVVRGDSVGGVITCVARNVRPGLSPVDFKLESELAKAMLSIPASNGPEI 302
Db  243 AEKMTAAIDTVRVRGDSIGGVVTCIARNVPRGLSPVDFKLESELAKAMLSIPASKGFEI 302
Qy  303 GSGFAGTDLTGSEHNDEFYMDKAGSVTRTRNRSRGVQGGISNVEIVHFKVAFKPTPSIGV 362
Db  303 GSGFAGTDTGSEHNDEFYMDAGNVRTRTRNRSRGVQGGISNGEIIYFKVAFKPTATIGK 362
Qy  363 KONTVSRERQNVELLARGHDPVAPRAVPVVSMAALVMDQLMAHVAOCFMPALNTAL 422
Db  363 KONTVSRHEDEVLLARGHDPVVPRAVPVVSMAALVMDQLMAHIAQCEMPFLNAL 422
Qy  423 QEPVGS 428
Db  423 QEPIGS 428

RESULT 7
US-10-425-114-57857
; Sequence 57857, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 57857
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73278F07_FLI.pep
```

```
US-10-425-114-57857

Query Match      81.4%; Score 1790.5; DB 15; Length 452;
Best Local Similarity 81.9%; Pred. No. 6.9e-169;
Matches 349; Conservative 31; Mismatches 43; Indels 3; Gaps 2;

Qy  4 VPKPOQVAHSRARLAPRAIGALLEFAPASSLSRFVAV-HRCRTARLEVKASGNTFGNYFQV 62
Db  17 VSQPPVSARASTRFLPRGIGALPESAP--TSLRLSVGRRRAASLEVKASGNVFGNYFQV 74
Qy  63 ATYGESHGGVGCVISGCPPIPLTEADLQVELDRRRPGQSRIITSTRKETDTCKILSGTH 122
Db  75 ATYGESHGGVGCVISGCPPIPLTEADLQVELDRRRPGQSRIITPRKETDTCKILSGTH 134
Qy  123 EGVTTGTPILVIVPNTDOIGSDHREIANVVRPSSHADATYDFKYGVRVAVQGGRRSGRTV 182
Db  135 DGMTTGTPIHVPNTDQGGDYSEMSKAYRPSHADATYDFKYGVRVAVQGGRRSSARETI 194
Qy  183 GRVAAGALPKKILKLCGLLEILSFVSKVHQVLPEDAVDYGSVTLQIESNIVRCPPDEY 242
Db  195 GRVAAGALAKKILKLSGVLEILAFVSKVHQVLPEDAVDYETVTLHIESNIVRCPPDEY 254
Qy  243 AEKMDAIDRVVRGDSVGGVITCVARNVRPGLSPVDFKLESELAKAMLSIPASNGPEI 302
Db  255 AEKMTAAIDTVRVRGDSIGGVVTCIARNVPRGLSPVDFKLESELAKAMLSIPASKGFEI 314
Qy  303 GSGFAGTDLTGSEHNDEFYMDKAGSVTRTRNRSRGVQGGISNVEIVHFKVAFKPTPSIGV 362
Db  315 GSGFAGTDTGSEHNDEFYMDAGNVRTRTRNRSRGVQGGISNGEIIYFKVAFKPTATIGK 374
Qy  363 KONTVSRERQNVELLARGHDPVAPRAVPVVSMAALVMDQLMAHVAOCFMPALNTAL 422
Db  375 KONTVSRHEDEVLLARGHDPVVPRAVPVVSMAALVMDQLMAHIAQCEMPFLNAL 434
Qy  423 QEPVGS 428
Db  435 QEPIGS 440

RESULT 8
US-10-437-963-201690
; Sequence 201690, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 201690
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_9703C.1.pep
US-10-437-963-201690

Query Match      78.4%; Score 1724; DB 16; Length 426;
Best Local Similarity 80.6%; Pred. No. 2.6e-162;
Matches 337; Conservative 31; Mismatches 32; Indels 18; Gaps 3;

Qy  15 ARLAPRAIG----ALFEFAPASSLSRFVAVHRCRTARLEVKASGNTFGNYFOVATYGESHG 70
Db  11 ARVLPRGGGGGFRAPPESAPA--SLRFSVGRRAARLEVKASANVFGNYFOVATYGESHG 68
```

QY 71 GGVGCVISGPPRIPLTEADLOVELDRRRPGQSRITSTRKETDTCKILSGTHEGVTTGP 130
DB 69 GGVGCVISGPPRIPLTEADLOVELDRRRPGQSRITSTRKETDTCKILSGTHEGVTTGP 128
QY 131 ILVIVNTQIGSDHREIANVYRPSHADATYDFKYGVRAVQGGSSGRKTVGRVAAGAL 190
DB 129 IHVVFVNTDQRGDYSMAKAYRPSHADATYDFKYGVRAVQGGSSGRKTVGRVAAGAL 188
QY 191 PKILKLCGLLEILSFVKVHVVLPEDAVDYGSVTLQIESNIVRCPPPEYAEKIMDAI 250
DB 189 AKILKLCGLLEILSFVKVHVVLPEDAVDYGSVTLQIESNIVRCPPPEYAEKIMDAI 248
QY 251 DRVVRGDSVGGVITCVARNVPRGLSPVFDKLESELAKAMLSIPASNGFEIGSGFAGTD 310
DB 249 DKVVRGDSVGGVITCVARNVPRGLSPVFDKLESELAKAMLSIPASNGFEIGSGFAGTD 308
QY 311 LTGSEHNDFFYMDKAGSVTRTRNSGVQGGISNIVHFKVAFKPTPSIGVKQNTVSR 370
DB 309 YTGSEHNDFFYMDKAGSVTRTRNSGVQGGISNIVHFKVAFKPTPSIGVKQNTVSR 356
QY 371 RONVELLARGHDPCCVAPRAVPVSMALVMDQLMAHVAOCFEMFALNTALQEPVGS 428
DB 357 HEDVELLARGHDPCCVAPRAVPVSMALVMDQLMAHVAOCFEMFALNTALQEPVGS 414

RESULT 9

US-10-268-611-39
; Sequence 39, Application US/10268611
; Publication No. US20030186352A1
; GENERAL INFORMATION:
; APPLICANT: MCLEOD, RIMA W.
; APPLICANT: ROBERTS, CRAIG W.
; APPLICANT: ROBERTS, FIONA
; APPLICANT: JOHNSON, JENNIFER J.
; APPLICANT: KIRISITS, MICHAEL
; APPLICANT: FERGUSON, DAVID
; APPLICANT: LYONS, RUSSELL
; APPLICANT: MUI, ERNEST
; APPLICANT: MACK, DOUG
; APPLICANT: SAMUEL, BENJAMIN
; APPLICANT: GORNICKI, PIOTR
; APPLICANT: ZUTHER, ELLEN
; TITLE OF INVENTION: APICOMPLEXAN CHORISMATE SYNTHASE SEQUENCES AND AN
; TITLE OF INVENTION: INHIBITOR OF THE SHIKIMATE PATHWAY
; FILE REFERENCE: 19338/93888
; CURRENT APPLICATION NUMBER: US/10/268,611
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: 09/103,331
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/329,269
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/346,036
; PRIOR FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 39
; TYPE: PRT
; LENGTH: 431
; ORGANISM: Solanum lycopersicum
US-10-268-611-39

Query Match 70.0%; Score 1541; DB 14; Length 431;
Best Local Similarity 72.7%; Pred. No. 4.4e-144;
Matches 287; Conservative 55; Mismatches 53; Indels 0; Gaps 0;
QY 32 SSSLPFAVHRCRTARLEVKASGNTFGNYFQVATYGESHGGVCGCVISGCPPIPLTEADL 91
DB 29 SSNLFPFTHRSQPKLEIQAAAGNTFGNYFRVTTFGESHGGVCGCIIIDGCPPLPSESMD 88
QY 92 QVELDRRRPGQSRITSTRKETDTCKILSGTHEGVTTGPILVIVNTDQIGSDHREIANV 151
DB 89 QVELDRRRPGQSRITSTRKETDTCKILSGTHEGVTTGPILVIVNTDQIGSDHREIANV 148
QY 92 QVELDRRRPGQSRITSTRKETDTCKILSGTHEGVTTGPILVIVNTDQIGSDHREIANV 151
DB 89 QVELDRRRPGQSRITSTRKETDTCKILSGTHEGVTTGPILVIVNTDQIGSDHREIANV 148

QY 152 YRPSHADATYDFKYGVRAVQGGSSGRKTVGRVAAGALPKILKLCGLLEILSFVKVH 211
DB 149 YRPSHADATYDFKYGVRAVQGGSSGRKTVGRVAAGALPKILKLCGLLEILSFVKVH 208
QY 212 QVVLPEDAVDYGSVTLQIESNIVRCPPPEYAEKIMDAIDRVRVRGDSVGGVITCVARNV 271
DB 209 NVVLPEDLVDNQIVTLQIESNIVRCPPPEYAEKIMDAIDRVRVRGDSVGGVITCVARNV 268
QY 272 PRGLSPVFDKLESELAKAMLSIPASNGFEIGSGFAGTDLTGSEHNDFFYMDKAGSVTR 331
DB 269 PRGLGTPTVPFDKLESELAKAMLSIPATKGFEGSGFAGTFTMTGSEHNDFFYMDKAGSVTR 328
QY 332 TNRSGVQGGISNIVHFKVAFKPTPSIGVKQNTVSRERQNVVELLARGHDPCCVAPRAV 391
DB 329 TNRSGVQGGISNIVHFKVAFKPTPSIGVKQNTVSRERQNVVELLARGHDPCCVAPRAV 388
QY 392 PVVEGMAALVMDQLMAHVAOCFEMFALNTALQEPV 426
DB 389 PVVEGMAALVMDQLMAHVAOCFEMFALNTALQEPV 423

RESULT 10

US-10-824-194-48
; Sequence 48, Application US/10824194
; Publication No. US20050142113A1
; GENERAL INFORMATION:
; APPLICANT: MCLEOD, RIMA W.
; APPLICANT: ROBERTS, CRAIG W.
; APPLICANT: ROBERTS, FIONA
; APPLICANT: JOHNSON, JENNIFER J.
; APPLICANT: KIRISITS, MICHAEL
; APPLICANT: FERGUSON, DAVID
; APPLICANT: LYONS, RUSSELL
; APPLICANT: MUI, ERNEST
; APPLICANT: MACK, DOUG
; APPLICANT: SAMUEL, BENJAMIN
; APPLICANT: GORNICKI, PIOTR
; APPLICANT: ZUTHER, ELLEN
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS, DIAGNOSTIC REAGENTS, AND VACCINES
; TITLE OF INVENTION: BASED ON UNIQUE APICOMPLEXAN PARASITE COMPONENTS
; FILE REFERENCE: 19338-90966
; CURRENT APPLICATION NUMBER: US/10/824,194
; CURRENT FILING DATE: 2004-04-14
; PRIOR APPLICATION NUMBER: US/09/631,594
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: PCT/US00/11478
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: PCT/US97/12497
; PRIOR FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 48
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Solanum lycopersicum
US-10-824-194-48

Query Match 70.0%; Score 1541; DB 18; Length 431;
Best Local Similarity 72.7%; Pred. No. 4.4e-144;
Matches 287; Conservative 55; Mismatches 53; Indels 0; Gaps 0;

QY 32 SSSLPFAVHRCRTARLEVKASGNTFGNYFQVATYGESHGGVCGCVISGCPPIPLTEADL 91
DB 29 SSNLFPFTHRSQPKLEIQAAAGNTFGNYFRVTTFGESHGGVCGCIIIDGCPPLPSESMD 88
QY 92 QVELDRRRPGQSRITSTRKETDTCKILSGTHEGVTTGPILVIVNTDQIGSDHREIANV 151
DB 89 QVELDRRRPGQSRITSTRKETDTCKILSGTHEGVTTGPILVIVNTDQIGSDHREIANV 148
QY 152 YRPSHADATYDFKYGVRAVQGGSSGRKTVGRVAAGALPKILKLCGLLEILSFVKVH 211
DB 149 YRPSHADATYDFKYGVRAVQGGSSGRKTVGRVAAGALPKILKLCGLLEILSFVKVH 208

212	QVULPEDAVDYGSVLTLEQIESINVRCPDPEAEKMIIDAIDRVVRGDSVGGVITCVARNV	271
Qy		
209	NWVLPEDLVDNQIVLTLEQIESINVRCPNPEAEKMIIGAIDYVRGDSVGGVITCIVRVN	268
Db		
272	PRGLGSPVDFDKLESELAKAMLSIPASNGFETGSGFAGTDLTGSEHNDPFYMDKAGSVRTR	331
Qy		
269	PRGLGTPVFDKLEAEKACMSLPATKGFEGSGFAGTFMTGSEHNDFFMDEHDOIRTK	328
Db		
332	TNRSGVGOGGISNVEIVHFVKAFKTPSGIVKQNTVSRRERQNVELLARGHPDPCVAPRAV	391
Qy		
329	TNRSGGIQGISNGEIIINRVAFKTSITARKQHTVSRDKHETELIARGHDPCCVVPRAV	388
Db		
392	PVTRSMALVLMDQLMAHVACQEMFALNTALQEPV	426
Qy		
389	PMVIRAMVALVLVDLMTQYAOCLMFPVNLTLQEPL	423
Db		

RESULT 11

```

US-10-268-611-63
; Sequence 63, Application US/10268611
; Publication No. US20030186352A1
; GENERAL INFORMATION:
; APPLICANT: MCLEOD, RIMA W.
; APPLICANT: ROBERTS, CRAIG W.
; APPLICANT: ROBERTS, JONNA
; APPLICANT: JOHNSON, JENNIFER J.
; APPLICANT: KIRISITS, MICHAEL
; APPLICANT: FERGUSON, DAVID
; APPLICANT: LYONS, RUSSELL
; APPLICANT: MUI, ERNEST
; APPLICANT: MACK, DOUG
; APPLICANT: SAMUEL, BENJAMIN
; APPLICANT: GORNICKI, PIOTR
; APPLICANT: ZUTHER, ELLEN
; TITLE OF INVENTION: APLICOMPLEXAN CHORISMATE SYNTHASE SEQUENCES AND AN
; TITLE OF INVENTION: INHIBITOR OF THE SHIKIMATE PATHWAY
; FILE REFERENCE: 19338/93888
; CURRENT APPLICATION NUMBER: US/10/268,611
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: 09/103,331
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/329,269
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/346,036
; PRIOR FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Solanum lycopersium
US-10-268-611-63

```

Query Match	69.4%	Score 1527;	DB 14;	Length 440;
Best Local Similarity	67.4%;	Pred. No. 1.1e-142;		
Matches 288;	Conservative 64;	Mismatches 73;	Indels 2;	Gaps 1;
Qy	4	VPKQQVAHSRAR--LAPRAITGALLEPAPASSSLRFAVHRCRTARLEVKASGNTFGNYFQ	61	
Db	5	VPTKQFVGASSSDIGSRLVSLQLPSKFSSSNFHLSPRPSQLKRLIQAGSITFGNYFR	64	
Qy	62	VATFGESHGGVGVCVISCPPRIPLTEADLOVELDRRRPQGSRTITSTKETTDTCKILSGT	121	
Db	65	VTTFGESHGGVGCTIIDCPRLPLSESDMQVELDRRRPQGSRTITTPRKETDTCKISSGT	124	
Qy	122	HEGVTGTPILVIVPNTDQIGSDHREIANVVRPSHADATYDFKGVRAVQGGSSSGRKT	181	
Db	125	ADGTTGSPIKVEPNTDQRNDYEMSLATRPSHADATYDFKGVRSVQGGSSSRARET	184	
Qy	182	VGRVAAGALPKKILKLCGLBEILGFVSKVQHVWLPEDAVDYGSVTLEQIESNIVRCPDPE	241	
Db	185	IGRVAAGAVAKKILKLSGAELVAVSOVHVPEDLIHONVTLQEIENIVRCPDPE	244	

Qy	242	YAEKMTAIDRVRVRGDSGGVITTCVARNVRGLGSPVFDKLESELAKAWLSIPASNGPE	301
Db	245	YAEKMTAAIDA VRVRGDSGGVVTCIVRNLPRLGTFVFDKLEAEELAKACMSLIPATKGE	304
Qy	302	IGSGFAGTDLTGSEHNDEFYMDKAGSVRTRTNRSGGVGQGISNVELTVHFVKVAFKPTPSIG	361
Db	305	FGSGFAGTFTWGSEHNDEFYMDHBGRI RTRTNRSGGIQQISNGEVINNRIGFKPTSTIS	364
Qy	362	VKQNTVSRSRQNVELLARGHDCVAPRAVPVSVESMAALVLMQLMAHVAAQCEMFALNTA	421
Db	365	RKQOTVTRDKHETELARGHDECVVPRAVPVMEANVALVLVDQLNAQYSCQMMFFPINPE	424
Qy	422	LQEPVGS	428
Db	425	LQEPLOS	431

```

RESULT 12
US-10-424-599-144653
; Sequence 144653, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 144653
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_101634C.1.pap
US-10-424-599-144653

```

RESULT 13
US-10-425-114-52276
; Sequence 52276, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 52276
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700829731_FLI.pep
US-10-425-114-52276

Query Match 68.7%; Score 1512; DB 15; Length 438;
Best Local Similarity 68.0%; Pred. No. 3.5e-141;
Matches 283; Conservative 62; Mismatches 63; Indels 8; Gaps 1;
QY 19 PRAIGALLEFPASS-----LRFVHRCRTARLEVKASGNTFGNYFOVATYGESHG 70
DB 12 PFSADALSASFSLNSDLGSLPPAYLRLSLRPLPKLHQAAGSTVGNHFRVTYGESHG 71
QY 71 GGVGVISCPRIPLTEADLQVELDRRPPGQSRTSTKTTCTKILSGTHEGVTGTP 130
DB 72 GGVGVICDCCPLPLSEADMVQLDRRPPGQSRTSTKTTCTKILSGTHEGVTGTP 131
QY 131 ILVIVPNTDQIGSDHREIANVRPSHADATYFKYGVRAVQGGSSGRKTVGRVAAGAL 190
DB 132 IHVSVENTDQRHDYSEMAVAYRPSHADATYDMKYGVRSVQGGSSARETIGRVASGAV 191
QY 191 PKKILKLCGLBILSVKSVHQPVEDAVDYGSTVLEQIESNIVRCPPPEYAEKMDAI 250
DB 192 AKKILKEPSTGTEILAVSVQHKVLPEDLIDHDTLTDQIESNIVRCPPPEYAEKMDAI 251
QY 251 DRVRGDSVGVVTCVARNVRGLGSPVFDKLESELAKAMLSIPASNGFETGSGFAGTD 310
DB 252 DAVVRGDSVGVVTCIVRNCPRGLGSPVFDKLESELAKAMLSIPATKGFQFGSGFAGTF 311
QY 311 LTGSHNDHDFYMDKAGSVTRTRNSGGVGGISNVEIVHFVAFKPTPSIGVKONTVSRRE 370
DB 312 LTGSHNDHDFYIDHGNTRTRNSGGIIGGINSGEIINMRVAFKPTSTIGKKQKTVTRD 371
QY 371 QNVELLARGHDPVAPRAVPVBSMAALVMDQLMAHVAQCEMFALNTALQEPV 426
DB 372 KKBTEFIARGHDPVCPVPRAPVPMVAMVAVLVDQLMAQYACQNLFPVNSDLQEPV 427

RESULT 14
US-10-696-616-28
; Sequence 28, Application US/10696616
; Publication No. US20040191850A1
; GENERAL INFORMATION:
; APPLICANT: Levin, Joshua
; APPLICANT: Tossberg, John
; APPLICANT: Zhou, Qing
; APPLICANT: McElver, John
; APPLICANT: Aux, George
; APPLICANT: Budziszewski, Greg
; APPLICANT: Thomas, Carla
; APPLICANT: Patton, David
; APPLICANT: Frye, Catherine

; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Essential for Plant Growth
; TITLE OF INVENTION: Development and Uses Thereof
; FILE REFERENCE: 70166USNP
; CURRENT APPLICATION NUMBER: US/10/696,616
; CURRENT FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: US 60/423,519
; PRIOR FILING DATE: 2002-11-04
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-696-616-28

Query Match 66.1%; Score 1454; DB 16; Length 436;
Best Local Similarity 67.7%; Pred. No. 2e-135;
Matches 268; Conservative 65; Mismatches 63; Indels 0; Gaps 0;
QY 31 ASSSLRFAVHRCRTARLEVKASGNTFGNYFOVATYGESHGCGVGVISGCPRIPLTEAD 90
DB 30 SSPAVQISLRTQTRKNFOIQATGSSYTHFRVSTFGSHGGVGCIIIDGCPRIPLTSSD 89
QY 91 LQVELDRRPPGQSRTSTKTTCTKILSGTHEGVTGTPILVIVPNTDQIGSDHREIAN 150
DB 90 LQFDLDRRPPGQSRTSTKTTCTKILSGTHEGVTGTPIHVFPVNTDQIGSDHREIAN 149
QY 151 VYRSHADATYDFKYGVRVAVQGGSSGRKTVGRVAAGALPKKILKLCGLBILSVKSV 210
DB 150 AYRPSHADATYDMKYGVRSVQGGSSARETIGRVAPGALAKKILKOPAGTEILAYVSVQ 209
QY 211 HQVLPEDAVDYGSTVLEQIESNIVRCPPPEYAEKMDAIDRVVRGDSVGVVTCVARN 270
DB 210 HHVLPBELVDHENTLEQIENNVRCPPPEYAEKMDAIDAVRTKGSVGVVTCIVRN 269
QY 271 VPRGLGSPVFDKLESELAKAMLSIPASNGFETGSGFAGTDLTGSHNDHDFYMDKAGSV 330
DB 270 APRGLGTPVFDKLESELAKAMLSIPATKGFQFGSGFAGTFLTGSHNDHDFYMDKAGSV 329
QY 331 RTNRSVGVGGISNVEIVHFVAFKPTPSIGVKONTVSRERONVELLARGHDPVAPRA 390
DB 330 RTNRSVGGIIGGINSGEIINMRVAFKPTSTIGRKONTVTRDKVETEMTARGHDPVAPRA 389
QY 391 VPVVSMAALVMDQLMAHVAQCEMFALNTALQEPV 426
DB 390 VPMVEMAVLVLDQLMAQYACQNLFPINPELQEPV 425

RESULT 15
US-10-677-179-12
; Sequence 12, Application US/10677179
; Publication No. US20040082050A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Penber, Stephen O.
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-A
; CURRENT APPLICATION NUMBER: US/10/677,179
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: US/09/743,207
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: PCT/US99/16353
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-677-179-12

Query Match	57.1%	Score 1257;	DB 15;	Length 307;
Best Local Similarity	81.9%;	Pred. No. 4.8e-116;		
Matches 245;	Conservative 22;	Mismatches 26;	Indels 6;	Gaps 2;
Qy	15	ARLAPRAIG----	ALLFAPASSLRFAVHRCRTARLEVKASGNTFGNYFQVATY	GSHG 70
Db	11	ARVLPRGGGGGFRAFPESAPA--	SLRFSVGRRRAARLEVKASANVFQVATY	GSHG 68
Qy	71	GGGCVVSGCPRPLPLEADLOVELDRRRPQOSRITSTRKETDCKILSGTHEGVT	TGP 130	
Db	69	GGGCVVSGCPRPLPLEADQVELDRRRPQOSRITTPRKETDCKILSGTHEGVT	TGP 128	
Qy	131	ILVTVPNTDQIGSDHREIANYRPSHADATYDFKYGVRVQGGGRSGRKT	VGRVAAGAL 190	
Db	129	IHFVPNTDQRGDGYSEMAKAYRPSHADATYDFKYGVRVQGGGRSARET	IGRVAAGAL 188	
Qy	191	PKKILKUKGLEILSFVSKVHONVLPEDAVDYGVTLEQTESNIVRCPDPEYAEK	MDAI 250	
Db	189	AKKILKUSGVEILAFVSKVHQVLPEDAVDITVTMEQTESNIVRCPDPEYAEK	MDAL 248	
Qy	251	DRVVRGDSVGGVITCVARNYPRGLGSPFDFKLESELAKAMLSIPASNGP	FIGSGFPAGT 309	
Db	249	DKVVRGDSIGGVVTCIARNYPRGIGSPFDFKLESELAKAMLSIPASKP	FIGSGFVPT 307	

Search completed: August 25, 2005, 19:35:16
Job time : 163 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 27, 2005, 02:58:29 ; Search time 4713 Seconds
(without alignments)

4410.628 Million cell updates/sec

Title: US-10-677-179-8

Perfect score: 2200

Sequence: 1 MTVPKPOQVAHSRLAPR.....VAQCEMFALTAQBPVGSF 429

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV-xlp

-Q/cn2_1/USPTO_spool_P/US10677179/runat_25082005_150240_24835/app query.fasta_1.583
-DB=GenEmbl -QFMT=fastap -SUFFIX=erge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10677179 @CN 1.1 5600 @runat_25082005_150240_24835 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

denEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	2200	100.0	1626	AR432943
2	1800	81.8	1607	AK099850
3	1797.5	81.7	1635	AR432940
4	1541	70.0	1802	LECHOSYNB
				Z21791 L.esculentu

5	1527	69.4	1701	8	LECHOSYNA
6	1486.5	67.6	1754	8	CSCHOSYN
7	1454	66.1	1311	6	CQ805346
8	1454	66.1	1311	8	BT015785
9	1454	66.1	1603	8	AY057519
10	1257	57.1	966	6	AR432945
C 11	1210.5	55.0	195583	8	AC134234
C 12	1195	54.3	110000	2	AP006487_1
C 13	1192.5	54.2	339650	1	AP003583
C 14	1191.5	54.2	1884	1	SRPAROC
C 15	1191.5	54.2	135551	1	D90906
C 16	1176.5	53.5	299350	1	AP005370
C 17	1066	48.5	1015	6	AR432942
18	1066	48.5	1057	8	BT009607
19	1047	47.6	349742	1	BX572090
20	1038	47.2	348764	1	BX569689
21	1025.5	46.6	306803	1	AE017161
22	1010	45.9	348071	1	BX572100
23	954.5	43.4	18062	8	AC084242
24	950.5	43.2	110000	1	CR522870_34
C 25	946.5	43.0	110000	8	CR382125_08
C 26	944	42.9	110000	8	CR380951_4
C 27	943	42.9	110000	8	AE016817_01
C 28	942	42.8	110000	8	CR382137_03
29	938	42.6	1876	8	SCVGL148W
30	938	42.6	3036	8	SCAROA
31	938	42.6	22803	8	SCV11228
C 32	930.5	42.3	110000	8	CR382130_22
33	930	42.3	1474	8	NCU25818
34	928.5	42.2	10157	1	AE011207
35	928.5	42.2	300693	1	AE017287
36	889.5	40.4	110000	1	AP006841_43
37	887	40.3	659	8	AB122084
38	876	39.8	59412	8	NCB7721
39	874.5	39.8	301731	1	AE016934
C 40	870	39.5	36677	2	AY114863
C 41	864	39.3	299180	1	AP005080
C 42	864	39.3	348526	1	CR378672
43	862	39.2	10923	1	AE013761
C 44	862	39.2	110000	1	BX936398_31
45	862	39.2	258050	1	AJ414153

ALIGNMENTS

RESULT 1
AR432943
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
ORIGIN

AR432943
Sequence 7 from patent US 6653531.
AR432943
AR432943.1 GI:40195584
Unknown.
Unclassified.
1 (bases 1 to 1626)
Cahoon,R.E. and Falco,S.C.
Chorismate synthase from plants
Patent: US 6653531-A 7 25-NOV-2003;
Location/Qualifiers
1. 1626
/organism="unknown"
/mol_type="genomic DNA"

DNA
1626 bp

linear PAT 18-DEC-2003

Alignment Scores:
Pred. No.: 1.47e-138
Score: 2200.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 6
Length: 1626
Matches: 429
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-10-677-179-8 (1-429) x AR432943 (1-1626)

Qy 1 MetThrValProLysProGlnValAlaHisSerArgAlaArgLeuAlaProArg 20
Db 79 ATGACGCGTGTCCCAAGCCACAGCAGGTGGCGCACTCACGGCAGCGCTCGCACCCGCG 138
Qy 21 AlalileGlyAlaLeuLeuGluPheAlaProAlaSerSerLeuArgPheAlaValHis 40
Db 139 GCGATCGGCGCTTCTCGAGTTTCCCGCAGCTCTCTCCCTCCGCTTCGCGTGCAC 198
Qy 41 ArgCysArgThrAlaArgLeuGluValLysAlaSerGlyAsnThrPheGlyAsnTyrPhe 60
Db 199 CGTCCGCACTGCTCGCTAGAGGTGAAGGCATCTGGAACACAGTTTGGAACTACTTTT 258
Qy 61 GlnValAlaThrTyrGlySerHisGlyGlyValGlyCysValIleSerGlyCys 80
Db 259 CAGGTTGCAACCTATGGTGAATCTCATGGGGTGGTGTGGTGTGTATCAGTGGTTGT 318
Qy 81 ProProArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArgArgPro 100
Db 319 CCACCTAGATTCCTACTCTAGGAGCAGACCTACAGTTGAATCGATCGAAGACGGCCC 378
Qy 101 GlyGlnSerArgIleThrSerThrArgLysGluThrAspThrCysLeuIleLeuSerGly 120
Db 379 GGACAGACGAGAATAACCTCCACAGGAAGAGAGACTGATACATGCAAAATTTCTGTACGG 438
Qy 121 ThrHisGluGlyValThrGlyThrProIleLeuValIleValProAsnThrAspGln 140
Db 439 ACACATGAAGGGGTGACTACTGGAACGCCAATCTTGTATTGTGCCAAACACAGATCAA 498
Qy 141 IleGlySerAspHisArgGluIleAlaAsnValTyrArgProSerHisAlaAsnAlaThr 160
Db 499 ATAGCAGTGATCACCGTGAATACCAATGTTACCGACCTTCTCATCGACAGCAACT 558
Qy 161 TyrAspPheLysTyrGlyValArgAlaValGlnGlyGlyArgSerSerGlyArgLys 180
Db 559 TATGACTTCAAGTACGGTGTAGAGCTGTACAGGGAGGTGGAGGTCTCTGGGCAGAAAA 618
Qy 181 ThrValGlyArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysLeuLysCysGly 200
Db 619 ACCGTTGGAAGGGTGGTGCAGGGGCCCTCCCAAGAAAAATTTCTTAAGCTCAAAATGTGGA 678
Qy 201 LeuGluIleLeuSerPheValSerLysValHisGlnValValLeuProGluAspAlaVal 220
Db 679 TTAGAGATCTGTGCTGTGTTGTTTCCAAAGTCATCAGTTGTGCTCCCAAGACACCGGTT 738
Qy 221 AspTyrGlySerValThrLeuGluGlnIleGluSerAsnIleValArgCysProAspPro 240
Db 739 GATTATGGTCTGTAACTTTGGAACAGATAGAGACCAATCGTTAGATGTCTCTGATCCA 798
Qy 241 GluTyrAlaGluLysMetIleAspAlaIleAspArgValArgValArgGlyAspSerVal 260
Db 799 GAGTACGACAGAGAAGATGATAGACCAATCGACAGATACGAGTTCGAGGGGATTCGGTC 858
Qy 261 GlyGlyValIleThrCysValAlaArgAsnValProArgGlyLeuGlySerProValPhe 280
Db 859 GGTGAGTGTATCATATGCTGCTAGAAACGTTCTTCGCGGCTCGGTTCTCTGTTGTTTC 918
Qy 281 AspLysLeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPhe 300
Db 919 GACAAGCTCGAATCCGAACTGGCAAAAGCTATGCTTCTATTCTTCGCGAGCAACGGGTTTC 978
Qy 301 GluIleGlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPhe 320
Db 979 GAGATTGGCAGCGGATTCGCGGGACCGCACTTGCAGGAGAGTGAACATATCGCTCGGTTGGTGCAGGGA 1038
Qy 321 TyrMetAspLysAlaGlySerValArgThrArgThrAsnArgSerGlyGlyValGlnGly 340
Db 1039 TATATGGATAGGCTGGAAAGTGTACGACACCGCACTAATCGCTCGGTTGGTGCAGGGA 1098
Qy 341 GlyIleSerAsnValGluIleValHisPheLysValAlaPheLysProThrProSerIle 360
Db 1099 GCGATATCGAATGTTGAGATTGTGCACTTCAAAGTTGCTTTTAAAGCCGACACCATCTATC 1158

Qy 361 GlyValLysGlnAsnThrValSerArgGluArgGlnAsnValGluLeuLeuAlaArgGly 380
Db 1159 GGGGTGAAACAGAACACCGGTCAAGGGAGCGTCAGAACCTTGGCTTCTAGCAAGAGGG 1218
Qy 381 ArgHisAspProCysValAlaProArgAlaValProValValGluSerMetAlaAlaLeu 400
Db 1219 CGCATGACCATGCGTGGCCCTCGAGCTGTCTGTGGTGGNAATCATGGCCGCTTG 1278
Qy 401 ValLeuMetAspGlnLeuMetAlaHisValAlaGlnCysGluMetPheAlaLeuAsnThr 420
Db 1279 GTCTCTATGGACCACTGATGGCGCACGTGGCTCAGTGCAGATGTTTCGGCTCAATACT 1338
Qy 421 AlaLeuGlnGluProValGlySerPhe 429
Db 1339 GCATTCAGAACCAAGTTGGCTCTTTTC 1365

RESULT 2
AK099850 1607 bp mRNA linear PLN 24-JUN-2003
LOCUS Oryza sativa (japonica cultivar-group) cDNA clone:J013110C10, full
DEFINITION insert sequence.
ACCESSION AK099850
VERSION AK099850.1 GI:32985059
KEYWORDS FLI CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Sukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1
AUTHORS The Rice Full-length cDNA Consortium, National Institute of
Agricultural Sciences Rice Full-length cDNA Project Team,
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiki, T., Ohmeda, E., Yahagi, W., Suzuki, K., Li, C.,
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group, Otono, Y., Murakami, K.,
Iida, Y., Sugeno, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Ikedara, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN;
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Otsu, N., Oca, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M. and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice

TITLE Science 301 (5631), 376-379 (2003)
JOURNAL MEDLINE
PUBMED 22752273
12869764
REFERENCE 2 (bases 1 to 1607)

AUTHORS Adachi, J., Aizawa, K., Akimura, T., Azakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
Hori, F., Hotta, I., Iida, Y., Ikeda, R., Inamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, O., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Namiki, T., Nariakawa, R., Niikura, J., Nishi, K., Nomura, K.,
Numasaki, R., Ohmeda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
Otsu, N., Oca, Y., Otono, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,

Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.
 Direct Submission
 Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)
 This clone is one of the 28K full-length cDNA clones from japonica rice.

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>
 NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ohta, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.
 PAIS Genome Sequencing & Analysis Group: Ohtsuki, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Mateubara, K. and Murakami, K.
 Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, P., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koyu, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Soabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, P., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES

Location/Qualifiers
 1..1607
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="J013110C10"

ORIGIN

Alignment Scores:
 Pred. No.: 9,99e-112 Length: 1607
 Score: 1800.00 Matches: 349
 Percent Similarity: 90.91% Conservative: 31
 Best Local Similarity: 83.49% Mismatches: 32
 Query Match: 81.82% Indels: 6
 DB: 8 Gaps: 2

US-10-677-179-8 (1-429) x AK099850 (1-1607)

QY 15 AlaArgLeuAlaProArgAlaIleGly-----AlaLeuLeuGluPheAlaPro 30
 DB 89 GCGCGGTCCTCCCGCGCGCGCGCGCGGGTTCGCGGCTTCCCGAGTCCGCCCG 148
 QY 31 AlaSerSerLeuArgPheAlaValHisArgCysArgThrAlaArgLeuGluVallys 50
 DB 149 GCT-----TCGCTCGGCTTCTCCGTCGCGCGCGCGCGCGCTCGCTAGAGGTGAAG 202
 QY 51 AlaSerGlyAenThrPheGlyAenThrPheGlnValAlaThrTyrGlyGluSerHisGly 70
 DB 203 GCGTCTGCAATGTATTTGGGAATCTTCCAGGTTCGAAGTTATGGAGAGTCTCATGGA 262
 QY 71 GlyGlyValGlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAsp 90
 DB 263 GCGCGTGTGGTGGTGAATCATGTGATGCGCCAGATGCCAGATCCCACTTACTGAACAGAT 322
 QY 91 LeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArglys 110
 ::

DB 323 ATGCAAGTAGAACTCGACGGGAGCGGCCAGCAGAGCAATAACCAACCCCAAGAAAG 382
 QY 111 GluThrAspThrCysLeuIleLeuSerGlyThrHisGluGlyValThrThrClyThrPro 130
 DB 393 GAGACTGACACTTGGAAAATTTCTTTCAGGGACACATGAAGGAATGACCACCTGGGACCCA 442
 QY 131 IleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsn 150
 DB 443 ATTCAATGTTTTTGTCCCGACACAGATCAGAGAGGGGGGTGATTACAGTGAATGCGTAAG 502
 QY 151 ValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaVal 170
 DB 503 GCCTACAGACCTTTCACATGCGAGTCAACTTATGACTTCAAAATACGGTGTGTAGACAGTG 562
 QY 171 GlnGlyGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeu 190
 DB 563 CAGGAGGTGGAAAGATCATCAGCAAGAGACCATTTGGAAGGGTGGCTGCAGAGAGCTCTT 622
 QY 191 ProLysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysVal 210
 DB 623 GCAAGAAAATTTCTTAAGCTCAATCTCGAGTAGAGATCTTGGCGTTTGTGTCCAAGGTG 682
 QY 211 HisGlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIle 230
 DB 683 CATCAAGTTGTACTACCAAGAGATGCCCTTGAATATGACACTGTAAACAATGGAACAGATA 742
 QY 231 GluSerAenIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIle 250
 DB 743 GAAAGCAACATGTTAGATGTCCTGATCCAGAAATATGCAGAGAAGATGATTGATGCAATC 802
 QY 251 AspArgValArgValArgGlyAspSerValGlyGlyValIleThrCysValAlaArgAsn 270
 DB 803 GATAAAGTACAGTTAGAGGTGATTTCGATTCGTGGTGGTGCACATGATGTCGAAGAAAT 862
 QY 271 ValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAla 290
 DB 863 GTTCTCTCGGGATTTGGCTCTCTCTGATTTTCACAAACTTGAGGCTGAATTTGCGGAAGCT 922
 QY 291 MetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAsp 310
 DB 923 ATGCTTTCTCTCTCTGCAAGCAAGGGGTTTCAGATGCGCAGTGGATTTGCGAGGTACTAC 982
 QY 311 LeuThrGlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThr 330
 DB 983 TACTCTGGAAGTGAACATAATGATGAGTTCTATATGGATGAGGCTGGAAATGTGAGAACA 1042
 QY 331 ArgThrAenArgSerGlyGlyValGlnGlyGlySerAenValGluIleValHisPhe 350
 DB 1043 AGAACTAATCGTTCAGCGGTGTGCAGGGAGGGATATCAAAATGGTGAATTTATATACCTC 1102
 QY 351 LysValAlaPheLysProThrProSerIleGlyValLysGlnAenThrValSerArgGlu 370
 DB 1103 AAGTAGCTTTCAAGCCAACCGCGACTATTGGGAAGAAACAACACTGTTTCAAGGGAG 1162
 QY 371 ArgGlnAenValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaProArgAla 390
 DB 1163 CATGAGGATGTGAATTTTAGCAAGGGCGCGCCAGCCCATGTTGTGTCTCCCTCGCGCT 1222
 QY 391 ValProValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisVal 410
 DB 1223 GTGCCGATGGTGGAGTCCATGGCGGCATTTAGTCTCATGGACACAGCTGATGGCGCACATT 1282
 QY 411 AlaGlnCysGluMetPheAlaLeuAenThrAlaLeuGlnGluProValGlySer 428
 DB 1283 GCTCAATGTGAGATGTTTCCACTGAACCTTGCTCTACAGAAGAACCAAGTTGGCTCT 1336

RESULT 3
 AR432940
 LOCUS AR432940 1635 bp DNA linear PAT 18-DEC-2003
 DEFINITION Sequence 1 from patent US 6653531.
 ACCESSION AR432940
 VERSION AR432940.1 GI:40195581
 KEYWORDS


```

source
1. .1802
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/strain="UC82B"
/db_xref="taxon:4081"
/clone="cLeCS2"
42. .1337
/EC_number="4.6.1.4"
/codon_start=1
/product="chorismate synthase 2"
/protein_id="CAA79854.1"
/db_xref="GI:410484"
/db_xref="GOA:Q42885"
/db_xref="UniProt/Swiss-Prot:Q42885"
/translation="NASSMLTKQFLGAPPSPGSGQSPKLCSSNLRPEPTRSOPKRL
ELQAGNTFEGYFRVTTHGESHGGVGLIDGPPRLPLSDDMVQLDRRPPGSRIR
TPRKETDTCIKSSADGLTTGSPIKVEVENTDQNDYSEMSLAYRSHADATYDF
KYGVRSGGGSSARETIGRVAAGAVAKILKYGSTELIAYSQHNVNVLPELDV
NQIVLEQIESNIVCPNPEYAEKMGIDIVVRGDSVGGVTCIVRNVRPLGTVP
FDKLAELAKACMSLPATKPEFGSGFAGTFTMGSEHNDPEFMDSHDQIRTKTNRSGG
LOGGISNGLINMRVAFKPTSTIAKQHTVSRDKHETELIARGHRDPCVVPRAVPMVE
ANVALVLDQMTQVQAQMLFPVNLTLQSPSTKSA"
42. .185
186. .1334
/product="chorismate synthase 2"
/EC_number="4.6.1.4"
1802

sig_peptide
mat_peptide

polyA_site
ORIGIN

Alignment Scores:
Score: 2,77e-94 Length: 1802
1541.00 Matches: 287
Percent Similarity: 86.58% Conservative: 55
Best Local Similarity: 72.66% Mismatches: 53
Query Match: 70.05% Indels: 0
DB: 8 Gaps: 0

US-10-677-179-8 (1-429) x LECHOSYNB (1-1802)
QY 32 SerSerSerLeuArgPheAlaValHisArgCysArgThrAlaArgLeuGluValHisAla 51
DB 126 TCTTCAAACTCCGATTCCTACTCACCAGATCATCCAGCCCAACAGCAGTACAGATCAAGCT 185
QY 52 SerGlyAsnThrPheGlyAsnThrPheGlnValAlaThrTyGlyGluSerHisGlyGly 71
DB 186 GCTGTAATACATTTGGAAATATTTCGCGGTACAAAGCTTTGGAGAAATCTCACGGTGT 245
QY 72 GlyValGlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAspLeu 91
DB 246 GGTGTCGGTGTGTAATTGATGATGATGTCCTCCCTCGACTCCCACTCTCAGAGTCGGATAG 305
QY 92 GlnValGluLeuAspArgArgArgProGlyGlnSerArgIleThrSerThrArgLysGlu 111
DB 306 CAAAGTGAACATTCAGCAGGAGGAGGCGAGTCAAGCCGAATTTACCACACCAAGAAAGAG 365
QY 112 ThrAspThrCysLeuIleLeuSerGlyThrHisGlyGlyValThrGlyThrProile 131
DB 366 ACTGATACTTGAATAATATCATCAGCACTGCTGATGTTTACACGGGATCTCAATC 425
QY 132 LeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnVal 151
DB 426 AAAGTTGAAGTACCCCACTACCCAGAGAGGAATGACTATAGTGAATGTCCTTGTCT 485
QY 152 TyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyGlyValArgAlaValGln 171
DB 486 TATAGGCCATCTCATGCTGATGCACTTATGACTTCAAGTATGGAGTTAGATCAGTACAG 545
QY 172 GlyGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuPro 191
DB 546 GGGGAGGTAGATCTCTCAGCAAGAAACCATCGGAGAGTGCCTGCTGGAGCTGTGCC 605
QY 192 LysLysIleLeuLysLysCysGlyLeuGluIleLeuSerPheValSerLysValHis 211

```

```

Db 606 AAGAAATTTCTCAAACTTTATTATTAGGAACTGAGATCCTTGTCTTATGTTTTCACAGTTCCAC 665
QY 212 GlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGlu 231
DB 666 AATGTTGTACTTCCAGAGGATTTAGTTGATTAACCCAGATTTGGACACTAGACAGATGAA 725
QY 232 SerAsnIleValArgCysProAspProGluTyAlaGluLysMetIleAspAlaIleAsp 251
DB 726 AGCAATATTGTTGTTGCCCGAACCTTGAATATGCAGAAAGATGATTTGGTGGCCATTCAT 785
QY 252 ArgValArgValArgGlyAspSerValGlyValIleThrCysValAlaArgAsnVal 271
DB 786 TATGTACGAGTGAGGCGGATTTCTGTTGGTGTGCTGTAACCTTGCACTTGTAGAAATGTT 845
QY 272 ProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMet 291
DB 846 CCAGAGGTCCTTGTGTACACCACTTTGATTAACCTTGAAGCTTGAGCTGGCTTAAAGCTTGC 905
QY 292 LeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeu 311
DB 906 ATGTCATTACAGCAACCAAGGGTTTGGAGTTTGGAGTGGCTTTCAGGTTACTTTTCATG 965
QY 312 ThrGlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThrArg 331
DB 966 ACTGCGAGTGAGCATTAATGATGATGTTCTTATGATGAACATGACCAAAATCAGGACAAA 1025
QY 332 ThrAsnArgSerGlyGlyValGlnGlyIleSerAsnValGluIleValHisPheLys 351
DB 1026 ACCAATAGATCTGGTGTATCCAGGAGGCATATCAAAATGGAGAAATTTAATATGAGA 1085
QY 352 ValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGluArg 371
DB 1086 GTAGCATTCAGCCCACTTCTACATTCAGAGGAGCAACATACATCTGTTAGTAGGATAAA 1145
QY 372 GlnAsnValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaProArgAlaVal 391
DB 1146 CATGAGACTGAATCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1205
QY 392 ProValValGluSerMetAlaLeuValLeuMetAspGlnLeuMetAlaHisValAla 411
DB 1206 CCGATGTTGGAAGCAATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1265
QY 412 GlnCysGluMetPheAlaLeuAsnThrAlaLeuGlnGluProVal 426
DB 1266 CAATGATGCTATTTCAGTCAATCTTACACTACAGGAACTTTG 1310

RESULT 5
LECHOSYNA 1701 bp mRNA linear PLN 06-FEB-1996
LOCUS L.esculentum chorismate synthase 1 precursor.
DEFINITION
ACCESSION Z21796
VERSION Z21796.1 GI:410481
KEYWORDS chorismate synthase 1.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 1701)
Gorlach,J., Schmid,J. and Amrhein,N.
Differential expression of tomato (Lycopersicon esculentum L.)
genes encoding shikimate pathway isoenzymes. II. Chorismate
synthase
Plant Mol. Biol. 23 (4), 707-716 (1993)
JOURNAL
MEDLINE 94072719
PUBMED 8251624
REFERENCE 2 (bases 1 to 1701)
Gorlach,J., Raesecke,H.R., Abel,G., Wehrli,R., Amrhein,N. and
Schmid,J.
Organ-specific differences in the ratio of alternatively spliced
chorismate synthase (LeCS2) transcripts in tomato
Plant J. 8 (3), 451-456 (1995)
JOURNAL
MEDLINE 96045551

```

```

PUBMED 7550381
REFERENCE 3 (bases 1 to 1701)
AUTHORS Goerlach,J.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-1993) Joern Goerlach, Institute of Plant
Sciences, Swiss Federal, Institute of Technology, Sonnegstr. 5,
Zuerich, CH-8092, Switzerland
FEATURES
source
1..1701
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="cLeCS1"
34..1356
/EC_number="4.6.1.4"
/product="chorismate synthase 1"
/protein_id="CAA79859.1"
/db_xref="GI:410482"
/db_xref="GOA:Q42884"
/db_xref="UniProt/Swiss-Prot:Q42884"
/translation="MASFVPTKQFVGASSSSDIGSRLVSLQLSFSSNFHLPSRP
SQLKLEIQAAAGTFGNFYRTTFEGSHGGVGCIIDGCPRLPLSDDMOVELDRR
PGSKRIITPRKETDTKISSGTADGLTTGSPIKVEVPNTDORGNDYSEMSLAYRPSHA
DATYDFKYGVRSVQGGRRSARETIGVAGAVAKKILKLYSGAEVLAYVQVHVNL
FEDLDHQNVIETQESNIVRCPDPPEYAEKRMIAIDAVRVGDSVGGVVTCIVRNLR
GLGTFPDKLBAELAKCMSPATKGFEGSFAGTFMTGSEHDFVMDHGHRTIR
NTRSGGIQGSINSEVINMRIGFKPTSTISRKQOTVTRDKHETELIARGHRHPCVPR
AVPMVEAMVALVLVDQLMAQYSQCMFPIINPELQELQSSPESAEVTL"
34..195
sig_peptide
mat_peptide
196..1353
/product="chorismate synthase 1"
/EC_number="4.6.1.4"
1701
polyA_site
ORIGIN
Alignment Scores:
Pred. No.: 2,248-93 Length: 1701
Score: 1527.00 Matches: 288
Percent Similarity: 82.44% Conservative: 64
Best Local Similarity: 67.45% Mismatches: 73
Query Match: 69.41% Indels: 2
DB: 8 Gaps: 1

US-10-677-179-8 (1-429) x LECHOSYNA (1-1701)
Qy 4 ValProLysProGlnGlnValAlaHisSerArgAlaArg-----LeuAlaProArgAla 21
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
46 GTACTACCAACAACTTTGTCGGAGCTTCGTCCTCCGATATCGGATCCTCTCGTCT 105
Qy 22 IleGlyAlaLeuLeuGluPheAlaProLaseSerSerLeuArgPheAlaValHisArg 41
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
106 GTATCTCTTCAGTACCTCTTAAGTTCTCATCTTCCAACTTTCATCTCCCATCTCGTCCA 165
Qy 42 CysArgThrAlaArgLeuGluValLysAlaSerGlyAsnThrPheGlyAsnThrPheGln 61
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
166 TCACAACCAACCGCTTAGAGATACAGGCTGCTGTGTAGTACATTGGAAATTTACTTCGCG 225
Qy 62 ValAlaThrTyrglySerHisGlyGlyValGlyCysValIleSerGlyCysPro 81
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
226 GTTACAACTTTGGAGATCTCATGCTGGTGGAGTTGTTGTTATTTATGATGATGATGCT 285
Qy 82 ProArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArgArgProGly 101
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
286 CCCCCTCTCCCACTTCTGAATCTCATATGACAGGTGGAACCTTGAACAGGAGGCGCAGGT 345
Qy 102 GlnSerArgIleThrSerThrArgLysGluThrAspThrCysValIleLeuSerGlyThr 121
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
346 CAAGCCGCATTTACCACCTTAGGAAGAGACTGACACTTGCAAAATTTTCATCAGGCACT 405
Qy 122 HisGluGlyValThrThrGlyThrProIleValIleValProAsnThrAspGlnIle 141
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

406 GCAGATGGCGTACTACTGAGTCTCCAATCAAGTTGAAGTCTCAACTGATGATCAGAGA 465

142 GlySerAspHisArgGluIleAlaAsnValTyArgProSerHisAlaAspAlaThrTyr 161

466 GGAATGACTACAGTGAATGTGCTGCTGTACAGGCCATCTCATGAGATGCACTTAT 525

162 AspPheLysTyrglyValArgAlaValGlnGlyGlyArgSerSerGlyArgLysThr 181

526 GACTTCAAGTATGAGTGAATCTGTACAGGGGGGTGATAGATCATCAGCAAGAGACT 585

182 ValGlyArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysLysCysGlyLeu 201

586 ATTGGGAGAGCTGCTGCTGGAGCAGTTGCTAAGAAAATTTCTCAAACTCTATTCTGGAGCT 645

202 GluIleLeuSerPheValSerLysValHisGlnValValLeuProGluAspAlaValAsp 221

646 GAGGTTCTTCTGTTATGTTCTCAAGTTACCAAGTTGTACTCTCTGAGGATTTGATGAT 705

222 TyrglySerValThrLeuGluGlnIleGluSerAsnIleValArgCysProAspProGlu 241

706 CATCAGAACGTGACTTTAGAGCAGATAGAAAGCAATATTGTCGATGCCGATCCGAA 765

242 TyrAlaGluLysMetIleAspAlaIleAspArgValArgGlyAspSerValGly 261

766 TATGCAGAGAAGATGATTGCTGCTATTGATGCTGACGAGTGAGAGGAGACTCTCTGTTGGT 825

262 GlyValIleThrCysValAlaArgAsnValProArgGlyLeuGlySerProValPheAsp 281

826 GGTGTGTAACTTGTGATTTGTAGAAATCTCCACGGGGTCTTGGTACACAGCTTTCGAT 885

282 LysLeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPheGlu 301

886 AAATTGAAGCTGAGCTAGCCAAAGCTTGTGATGCTACTCTCGCACAAAGGGTTTGGAG 945

302 IleGlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPheTyr 321

946 TTTGGCAGTGGCTTTGCGAGCACGTTTCATGACTGTTAGCGAGCATATGATGAATCTAT 1005

322 MetAspLysAlaGlySerValArgThrArgThrAsnArgSerGlyGlyValGlnGlyGly 341

1006 ATGGATGAGCATGCCCGAATCAGGACAAAGAACCAACAGATCTGTGTGAATCCAGGGCGGT 1065

342 IleSerAsnValGluIleValHisPheLysValAlaPheLysProThrProSerIleGly 361

1066 ATATCAATGGAGAGTATCAATATGAGATAGTGTTCAGGCCAATCTCACTATTCTTCT 1125

362 ValLysGlnAsnThrValSerArgGluArgGlnAsnValGluLeuLeuAlaArgGlyArg 381

1126 AGGAAGCAGCAAACTGTGACAGAGACAAACACAGAAACAGAACTCATCGTAGGGTCGC 1185

382 HisAspProCysValAlaProArgAlaValProValValGluSerMetAlaAlaLeuVal 401

1186 CATGATCTTGTGTGGTTCCTCCGAGCTGTTCCAATGTTGAAGCAATGGTCGCTTGGTG 1245

402 LeuMetAspGlnLeuMetAlaHisValAlaGlnCysGluMetPheAlaLeuAsnThrAla 421

1246 CTGTTGATCAGTTAATGGCTCAGTATTCAGTGATGATGTTCCCAATCAATCTCTGAA 1305

422 LeuGlnGluProValGlySer 428

1306 CTACAGGAACCTTTGAGTCA 1326

RESULT 6

CSCHOSYN 1754 bp mRNA linear PLN 03-DEC-1991

LOCUS C.sempervirens mRNA for chorismate synthase precursor.

DEFINITION X60544

ACCESSION X60544.1 GI:18255

VERSION

KEYWORDS chorismate synthase.

SOURCE Corydalis sempervirens (pink corydalis)

ORGANISM Corydalis sempervirens

Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;

Fumariaceae; Corydalis.
 1 (bases 1 to 1754)
 Schaller, A., Schmid, J., Leibinger, U. and Amrhein, N.
 Molecular cloning and analysis of a cDNA coding for chorismate synthase from the higher plant Corydalis sempervirens Pers
 J. Biol. Chem. 266 (32), 21434-21438 (1991)
 92042037
 1718979
 2 (bases 1 to 1754)
 Amrhein, N.
 Direct Submission
 Submitted (18-JUL-1991) N. Amrhein, Swiss Federal Institute of Technology, Institute of Plant Sciences, Sonnegstr 5, Zuerich 8092, Switzerland
 Location/Qualifiers
 1..1754
 /organism="Corydalis sempervirens"
 /mol_type="mRNA"
 /db_xref="taxon:3464"
 1..1754
 /evidence=experimental
 165..1508
 /BC_number="4.6.1.4"
 /codon_start=1
 /product="chorismate synthase"
 /protein_id="CAA43034.1"
 /db_xref="GI:18256"
 /db_xref="GOA:P27793"
 /db_xref="UniProt/Swiss-Prot:P27793"
 /translation="MASSLSITKPFISGRRRTTDSGWSYFQTSDLRLQNSQVQIS VRRQAPLKLVAQSGSFGVQVTTTGESHGSGVGVDCGPPRFPFISEADIIQSDL DRRPQQRITTPRKETDKIYGVADFTTGSPIHISVPNTDORNDYSEMAKAYR PSHADATYDFKVSQSGSSARETIGRVAAGALAKKILKAVAGTEVLAVUSOAH KVPLEGLVDHETLSLEQIESNIVRCPSSEYAEKMLAIDAVRVKDSVGVVTCIMR NVPRGLSVDFKLAEALAKAMSLPATKGFSGFTFLTGSEHNDFFYDNGR IYTRNRSQIGGSLNGSIIINRIAFKPTSTIGKQNTVTREREIEILIRGRHDP C VYPRAPVMEANVALVLLDQLMLOAQGNLFSINPALOEPLSETVSSAAASIQGV"
 transit peptide 165..335
 mat_peptide 336..1505
 /product="chorismate synthase"
 /BC_number="4.6.1.4"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1 21e-90 Length: 1754
 Score: 1486.50 Matches: 280
 Percent Similarity: 85.43% Conservative: 60
 Best Local Similarity: 70.35% Mismatches: 55
 Query Match: 67.57% Indels: 3
 DB: 8 Gaps: 2
 US-10-677-179-8 (1-429) x CSCHOSYN (1-1754)
 QY 31 AlaSerSerSerLeuArgPheAlaValHisArgCysArgThrAla-----ArgLeuGlu 48
 Db 273 TCDAATCAATCGTCAGATCTCTGTCGACGA---CAAACCGCCCTTGAACACTAGTT 329
 QY 49 ValLysAlaSerGlyAsnThrPheGlyAenThrPheGlnValAlaThrTrpGlyGluSer 68
 Db 330 GTACAGGATCGCGGAGTTCAVTTGGGAAAGTCTTCCAAAGTTACACATATGGAGATCT 389
 QY 69 HisGlyGlyValGlyValGlyValIleSerGlyCysProArgIleProLeuThrGlu 88
 Db 390 CACGAGAGGAGTGTGGTGTGTGTCATGTATGATGTCCTCCACGGTTTCCCATCTCCGAA 449
 QY 89 AlaAspLeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThr 108
 Db 450 GCTGATATACATCAGATCTTCACAGAGGAGCGCAGGTACAGCGCATAACTACCCCA 509
 QY 109 ArgLysGluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrGly 128
 Db 510 AGAAAGAAACTGATACATGCAAAATATATTTCTGGAGTGTCTGATGGATTTCACAACTGGG 569
 QY 129 ThrProIleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIle 148
 Db 570 TCTCCAAATTCACATCAGTGTGCTTACACTGATCAGCGAGGAATGATTTACAGTGAATG 629
 QY 149 AlaAsnValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArg 168
 Db 630 GC AAAAGCTTATAGGCTTCTCATGCTGATGCTACTATGACTTCAAGTATGGTGTGAGA 689
 QY 169 AlaValGlnGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGly 188
 Db 690 TCAGTGCAGGGTGGGGAGATCTTCAGCAAGAGAAACCATTTGGGAGAGTTTCAGCTGGA 749
 QY 189 AlaLeuProLysIleLeuLysCysGlyLeuGluIleLeuSerPheValSer 208
 Db 750 GCTCTTGCAGAGATTTCTCAAGGCTTATGCAGGACTGAGGTCTAGCATATGTCTCT 809
 QY 209 LysValHisGlnValIleValProGluAspAlaValAspTyrGlySerValThrLeuGlu 228
 Db 810 CAAAGCCCAAAAGTTGTTCTCCAGAGGGTTTGGTTGATCATGAGACTCTATCACTCGAA 869
 QY 229 GlnIleGluSerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAsp 248
 Db 870 CAGATTGAAGTAACATCGTTAGTGCACAGATTCAGAAATACCCGAGAGATGATGCT 929
 QY 249 AlaIleAspArgValArgValArgCysProAspValGlyGlyValIleThrCysValAla 268
 Db 930 GCTATTGATGCTGCTCCGAGTGAAGGAGACTCTGTTGTTGGAGTTGTGACATGATATG 989
 QY 269 ArgAsnValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAla 288
 Db 990 AGGAACGTCCTCCAGCTGGCTTGGTTCACCAAGTATTTGATAAATGAGCTGAGCTTGC 1049
 QY 289 LysAlaMetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGly 308
 Db 1050 AAAGCGTGCATCTCATACCAACAAAGGGGTTTGAATTTGGGAGTGGATTTTCAGGC 1109
 QY 309 ThrAspLeuThrGlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerVal 328
 Db 1110 ACATTTCTGACTGTTAGTGAACATATATGATGTTCTACACAGACAGAGATGGAAGATC 1169
 QY 329 ArgThrArgThrAsnArgSerGlyValGlnGlyGlyIleSerAsnValGluIleVal 348
 Db 1170 AGGACTAGAACAAATCGATCTGTTGGAATACAGGGTGTATATATCAACGAGAGATCATA 1229
 QY 349 HisPheLysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSer 368
 Db 1230 AACATGAGATAGCTTTTCAGCCGACATCTACAATTTGGAAAGAGACAGATACTGTGACC 1289
 QY 369 ArgGluArgGlnAsnValGluLeuAlaArgGlyArgHisAspProCysValAlaPro 388
 Db 1290 AGAGAACGTGAAGATTTGAATCTATGACGTGTGTCGACGACCCCTTGTGTACTT 1349
 QY 389 ArgAlaValProValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAla 408
 Db 1350 CGAGCTGTTCCATGTTGGAAGCTATGTATGTCCTGCTGCTCTGATCACTGATGCTT 1409
 QY 409 HisValAlaGlnCysGluMetPheAlaLeuAsnThrAlaLeuGlnGluProVal 426
 Db 1410 CAACATGCTCAGGCAACCTTTTCTCCATCAATCTGCTTACAGAGACCTTTG 1463
 RESULT 7
 CQ805346 1311 bp DNA linear PAT 10-MAY-2004
 LOCUS Sequence 1757 from Patent WO2004/035798.
 DEFINITION CQ805346
 ACCESSION CQ805346
 VERSION CQ805346.1 GI:47111307
 KEYWORDS Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
 1
 REFERENCE

AUTHORS Inze, D., de Veylder, L. and Vlieghe, K.
IDENTIFICATION OF NOVEL E2F TARGET GENES AND USE THEREOF
JOURNAL Patent: WO 2004035798-A 1757 29-APR-2004,
CropDesign N.V. (BE)

FEATURES

Location/Qualifiers
1..1311
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN

Alignment Scores:
Pred. No.: 1 27e-88 Length: 1311
Score: 1454.00 Matches: 268
Percent Similarity: 84.09% Conservative: 65
Best Local Similarity: 67.68% Mismatches: 63
Query Match: 66.09% Indels: 0
DB: 6 Gaps: 0

US-10-677-179-8 (1-429) x CQ805346 (1-1311)

Qy 31 AlaSerSerLeuArgPheAlaValHisArgCysArgThrAlaArgLeuGluVallys 50
Db 88 TCTTCTCCCGCGTTTCAGATCTCTCTCCGTACCCAAACCCAGGAAGAACTTCCAGATACA 147
Qy 51 AlaSerGlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGly 70
Db 148 GCTACTGGAAGTTCATATGGGACTCATTTTCGAGTTTCAACTTTTGGAGAAATCACTGGA 207

Qy 71 GlyGlyValGlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAsp 90
Db 208 GGAGAGTGGTTGTATCATGATGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 267

Qy 91 LeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArglys 110
Db 268 TTGCAATTCGATCTCGATAGAGGAGGCTGCTCAGACGAGCATCACAACTCTTAGAAA 327

Qy 111 GluThrAspThrCysLeuIleLeuSerGlyThrHisGluGlyValThrThrGlyThrPro 130
Db 328 GAGACTGATACTTGGCGGATTCGCTCGAGTCTCTGAAGGAATGACACAGGAAACACT 387

Qy 131 IleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsn 150
Db 398 ATCCATGTTGTTGATACCAACACAGATCAGAGAGCACTGATTACAGTGAATGTCGTT 447

Qy 151 ValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaVal 170
Db 448 GCCTATAGACCATCGCATGCTGATGCAACTTATGACATGAAATGTTGTCAGATCAGTG 507

Qy 171 GlnGlyGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyValLeu 190
Db 508 CAGGTGGAGGAAGATCTTCAGCTAGAGAGCACTTGAAGAGTTCCTCGGAGCTTTG 567

Qy 191 ProLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysVal 210
Db 568 GCCAAGAAATTTTGAAGCAATTCAGGAAGTTCAGATCTTGCCTATGTTCTCGCAAGTT 627

Qy 211 HisGlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIle 230
Db 628 CACCATGTTGATCTCCAGAAATTTGGTAGACACGAGAAATTTAACTCGAACAGATA 687

Qy 231 GluSerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIle 250
Db 688 GAAATAACATTTGTGATGTCCTTAATCCCGAGTATCGGAAAGATGATAGTCCGAT 747

Qy 251 AspArgValArgValArgLysSerValGlyGlyValIleThrCysValAlaAlaArgAsn 270
Db 748 GATGCTGTGACGAGCAAAAGGAACTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 807

Qy 271 ValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAla 290
Db 808 GCTCCACGTTGGGTACACCGGTTTTCGATAAACTTGAAGCAGAACTCGCAAAAGCT 867

Qy 291 MetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAsp 310
Db 868 TGTATGTCGTACTGCACTGAAGGATTTGAGTTTGAAGCGCTTTGACGGTACCTTT 927

Qy 311 LeuThrGlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThr 330
Db 928 TTGACTGCTTCTGAACACATGATGAGTTCATACCGATGAAATGGAAGATACGTACC 987

Qy 331 ArgThrAsnArgSerGlyGlyValGlnGlyGlySerAsnValGluIleValHisPhe 350
Db 988 AGAACCAACCGATCTGCTGGAATTCAGGAGGATCTCAAAATGCTGTAATTAATTAACATG 1047

Qy 351 LysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGlu 370
Db 1048 AGATAGCTTCAAGCCACATCAACATTTGGAAGGAAGCAAAATACGTTAACCGAGAC 1107

Qy 371 ArgGlnAsnValGluLeuAlaArgGlyArgHisAspProCysValAlaAlaProArgAla 390
Db 1108 AAGGTAGAACCGAATGATTCGCGTGGTTCGTCATGATCTCTGTTGTTTCCAGGACT 1167

Qy 391 ValProValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisVal 410
Db 1168 GTGCCAATGTTGGAAGCAATGCTGCTCTGTTCTGTCGATCAATTCATGTCGCAATAC 1227

Qy 411 AlaGlnCysGluMetPheAlaLeuAsnThrAlaLeuGlnGluProVal 426
Db 1228 GCACATGCCATTTGTTTCCAAATAATCCAGATTGTCAGGAACCTCTC 1275

RESULT 8

BT015785 1311 bp mRNA linear PLN 02-OCT-2004
LOCUS Arabidopsis thaliana Atlg48850 mRNA, complete cds.
DEFINITION BT015785
VERSION BT015785.1 GI:53749179
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

1 (bases 1 to 1311)

Shinn, P., Chen, H., Cheuk, R., Kim, C.J. and Ecker, J.R.

Arabidopsis ORF clones

Unpublished

2 (bases 1 to 1311)

Shinn, P., Chen, H., Cheuk, R., Kim, C.J. and Ecker, J.R.

Direct Submission

Submitted (02-OCT-2004) Salk Institute Genomic Analysis Laboratory
(SIGnAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA

Location/Qualifiers

1..1311
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clones="U67693"
/ecotype="Columbia"
/note="This clone is in pUNI 51"

1..1311

/note="unknown protein"

/codon_start=1

/product="Atlg48850"

/protein_id="AAU90075.1"

/db_xref="GI:53749180"

/translation="MASSSLTSKILGSTKLGSSSLPSELRRLSSPAVOISLRTOTRK
NFQIQATGSSYGTFRFVSTGESHGGVGLIDGCPRIPLTSDLDLDRREGOS
RITTRKETDTCRISSGVSGMTTGTPIHVFPNTDORGLDYSMSVAYRPSHADATY
DMKYGVSVQGGKSSARETIGRVAPGALKAILKQFAGTEIVGVVQVHRVLPPEL
VDHENLTLEQIENNPVPAERKMAIAIDAVRTKSGVGGVVTCTVRNAPRLGT
PVFDKLAELAKACMSLPATKGFSGFAGTFLTGLEHNDFTLGTENGRIRTRTNR
GGIQGGISNGEIIINMRVAFKPTSTIGRKQNTVTRDKVETEMIAARGHDPCCVVPRAVPM

VEAMVALVLDQLMAQYAQCHLFFINPELOEPLOIEPONATAL"

QY	331	ArgThrAsnArgSerGlyValGlyValGlnGlyIleSerAsnValGluIleValHisPhe	350
QY	331	ArgThrAsnArgSerGlyValGlyValGlnGlyIleSerAsnValGluIleValHisPhe	350
Db	988	AGAACCACCGATCTGTGGAAATTCAGGAGGGATCTCAATGGTGAATAATAAATG	1047
QY	351	LysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGlu	370
QY	351	LysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGlu	370
Db	1048	AGAGTAGCTTCAAGCCACATCAACAATTTGGAGGAAGCAAAATACGGTAAACCAGAGAC	1107
QY	371	ArgGlnAsnValGluLeuLeuAlaAAGGlyArgHisAspProCysValAlaProArgAla	390
QY	371	ArgGlnAsnValGluLeuLeuAlaAAGGlyArgHisAspProCysValAlaProArgAla	390
Db	1108	AAGTAGAACCAGAAATGATGGCGTGGTGCATCATCTCTGTGTCTTCACGAGCT	1167
QY	391	ValProValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisVal	410
QY	391	ValProValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisVal	410
Db	1168	GTGCCAATGTGGGAAGCAATGGTGGCTCTAGTTCTTGTGTGATCAATTGATGGCGCAATAC	1227
QY	411	AlaGlnCysGluMetPheAlaLeuAsnThrAlaLeuGlnGluProVal	426
QY	411	AlaGlnCysGluMetPheAlaLeuAsnThrAlaLeuGlnGluProVal	426
Db	1228	GCAAAATGCCAATTTGTTTCCAAATAAATCCAGAGTTGCAGAGAACTCTTC	1275
Db	1228	GCAAAATGCCAATTTGTTTCCAAATAAATCCAGAGTTGCAGAGAACTCTTC	1275

RESULT 9

AV057519	LOCUS	AV057519	1603 bp	mRNA	linear	PLN 07-OCT-2001
	DEFINITION	Arabidopsis thaliana At1g48850/m24p22_3				
	ACCESSION	AV057519				
	VERSION	AV057519.1				
	KEYWORDS	FLI CDNA.	GI:15982823			
	SOURCE	Arabidopsis thaliana (thale cress)				
	ORGANISM	Arabidopsis thaliana				

ORGANISM
Arabisopsis thaliana (Lamarck 1800)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 1603)

REFERENCES

AUTHORS

Cheuk, R., Chen, H., Kim, C. J., Koessena, E., Meyers, M. C., Banh, J., Bower, L., Carninci, P., Dale, J. M., Goldsmith, A. D., Hayashizaki, Y., Ikehida, J., Jiang, P. X., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J. M., Lin, J., Liu, S. X., Miranda, M. K., Narusaka, M., Nguyen, M., Onodera, C. S., Palm, C. J., Pham, P. K., Quach, H. L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C. C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R. W., Theologis, A. and Ecker, J. R.

TITLE	JOURNAL	REFERENCE	AUTHORS
Arabidopsis CDNA clones	Unpublished		
2 (bases 1 to 1603)			
Bower, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.			
Chen, R., Chen, H., Kim, C.J., Koessena, E., Meyers, M.C., Banh, J., Bower, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.			

TITLE

JOURNAL: Submitted (18-SEP-2001) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA

COMMENT

NIH Biomedical Research Center (1996). Members carried out the collection and clustering of RAPL cDNAs : 'RIKEN Arabidopsis Full-length cDNA' : Seki, M., Narusaka, M., Ishida, J., Stou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Havaishizaki, Y. and Shinozaki, K.

The Salk, Stanford, POEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Cheuk, R., Chen, H., Kim, C. J., Koeseema, E., Meyers, M. C., Shinn, P., Bahh, J., Bowser, L., Dale, J. M., Goldsmith, A. D., Liang, P. X., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J. M., Lin, J., Liu, S. X., Miranda, M., Nguyen, M., Onodera, C. S., Palm, C. J., Pham, P. K., Quach, H. B., Southwick, A., Tang, C. C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R. W., Theologis, A., and Ecker, J. R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to

ORIGIN

Alignment Scores:
 Pred. No.: 1,27e-88 Length: 1311
 Score: 1454.00 Matches: 268
 Percent Similarity: 84.09% Conservative: 65
 Best Local Similarity: 67.68% Mismatches: 63
 Query Match: 66.09% Indels: 0
 DB: 8 Gaps: 0

US-10-677-179-8 (1-429) x BT015785 (1-1311)

Qy 31 AlaSerSerSerLeuArgPheAlaValHisArgCysArgThrAlaArgLeuGluVallys 50
 88 TCTTCTCCCGCGTTCAGATCTCTCCGTACCCAAACCAGGAAGAACTTCCAGATACAA 147
 Qy 51 AlaSerGlyAsnThrPheGlyAsnThrPheGlnValAlaThrTyArgGlyGluSerHisGly 70
 148 GCTACTGGGAAGTTCCATATGGGACTCATTTTCAGTTTCAACTTTTCGGGAATCACATGGA 207
 Qy 71 GlyGlyValGlyCysValIleSerGlyCysProProArgIleProLeuThrGluAlaAsp 90
 208 GGAGGAGTGGTGTATCATTCATGGTGTCTCTCTCGTATTCCTACTTACTGAATCTGAT 267
 Qy 91 LeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLys 110
 268 TTGCAATTCGATCTCGATAGAAGGAGGCCCTGGTCAGAGCAGGATCAACAACCTCTAGAAA 327
 Qy 111 GluThrAspThrCysValIleLeuSerGlyThrHisGluGlyValThrThrGlyThrPro 130
 328 GAGACTGATACTTGGCGGATATCGTCTGGAGTCTCTGAAGGAATGACGACGAGAAACACCT 387
 Qy 131 IleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsn 150
 388 ATCCATGTGTTTGTACCAACACACATCAGAGGAGCTTGATTTACAGTGAATGTGCGTT 447
 Qy 151 ValTyArgProSerHisAlaAspAlaThrTyArgPheLysTyArgValArgAlaVal 170
 448 GCCTATAGACCATCGATGCTGATGCACTTATGACATGAGTATGGTGTCTCAGATCAGTG 507
 Qy 171 GlnGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeu 190
 508 CAGGCTGGAGGAAGATCTTCAGTACAGAGAGACATTTGGAAGAGTTGCTCCTCGAGCTTG 567
 Qy 191 ProLysValIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysVal 210
 568 GCGAAGAAATTTTGAAGCAATTTGCAGAACTGCAGAACTGAGATTTCTGCTCATGTCTCG 627
 Qy 211 HisGlnValValLeuProGluAspAlaValAspTyArgLysValThrLeuGluGlnIle 230
 628 CACCATGTTGTACTTCCAGAGAATTTGTAGACCACGAGAATTTAACTCTCGACAGATA 687
 Qy 231 GluSerAsnIleValArgCysProAspProGluTyArgAlaGluLysMetIleAspAlaIle 250
 688 GAAATAACAATTTGTCAGATGCCCTAATCCGAGTATCGGAAAGATGATAGTCTGCGATT 747
 Qy 251 AspArgValArgValArgGlyAspSerValGlyValIleThrCysValAlaArgAsn 270
 748 GATGCTGTCCAGACAAAAGGGAACCTCTGTGTGTGTGTGTGTGACCTGCACTTGTTCGGAT 807
 Qy 271 ValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAla 290
 808 GCTCCAGCTGGGCTTGGTACACCGGTTTTCGATAAATTTGAAGCAGAACTGCGAANAAGCT 867
 Qy 291 MetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAsp 310
 868 TGTATGTCGCTACCTGCAACTAAGGGAATTTGAGTTTGGAAAGCGCTTTGCAGGTACCTTT 927
 Qy 311 LeuThrGlySerGluHisAsnAspGluPheTyMetAspLysAlaGlySerValArgThr 330
 928 TTGACTGCTTGTGAACAACATGATCAGTCTTATACCGATGAAATGGAAGATGATGCTACC 987

this work. Shinozaki, K. (RIKEN GSC) and Becker, J.R. (ssp/salk) contributed equally to this work as PIs.

```
FEATURES
    source
        1..1603
            /organism="Arabidopsis thaliana"
            /mol_type="mRNA"
            /db_xref="taxon:3702"
            /chromosome="1"
            /clone="RAPL09-47-O14 (R14512)"
            /ecotype="Columbia"
            1..122
            123..1433
                /note="putative chorismate synthase"
                /codon_start=1
                /product="At1g48850/T24P22_3"
                /protein_id="AAI09759.1"
                /db_xref="GI:15982824"
                /translation="MASSSLTSKSLGSLKSLSSLPSELRLRSLSPAVQISLRITQTRK
                NFQIQATSSYTHPRVSTFGESHGVCIIIDGPPRIPLTSDLOPLDRRPGOS
                RITPRKETDTCRISSGVSEGMWTTCTPIHVFPNTDORGLDYSMSVAYRSHADATY
                DMKYGVRVQSGGRSSARETIGRVAPGALAKKILKQFAGTBILAYVQVHHVVLPEEL
                VDHEMLTUEQIENNICVRCNPFEYAEKMIADAVRTKNSVGGVVTCLVRNAPRLGLT
                PVFDLEAEAKASLSPATKGFPGSGFAGTFLTGLEHNDFFYTDENGRITRTNRS
                GGIQGGISNGBIINNRVAFKPTSTIGRKNQTVTRDKVETEMIAARGHDPCCVPRVPM
                VEAMVALVLVQLMAQYAQCHLFPINPELOPQLQIEQPQNATAL"
            1434..1603

3' UTR
ORIGIN
Alignment Scores:
Pred. No.:      1,64e-88      Length:      1603
Score:          1454.00      Matches:    268
Percent Similarity: 84.09%      Conservative: 65
Best Local Similarity: 67.68%      Mismatches: 63
Query Match:      66.09%      Indels:     0
DB:               8          Gaps:         0

US-10-677-179-8 (1-429) x AY057519 (1-1603)

Qy  31 AlaSerSerLeuArgPheAlaValHisArgCysArgThrAlaArgLeuGluVallys 50
Db  210 TCTTCTCCCGCGGTCAGATCTCTCCGTCACCCAAACAGGAAGAACTTCAGATACAA 269
Qy  51 AlaSerGlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGly 70
Db  270 GCTACTCGAGAGTTCATATGGGACTCATTTTCAGTTTCAACTTTTGGAGATCAGATGA 329
Qy  71 GlyGlyValGlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAsp 90
Db  330 GGAGGAGTTGGTTGTATCATTTGATGTTGTCCTCGTATTCACCTTACTGAAATCTGAT 389
Qy  91 LeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArglys 110
Db  390 TTGCAATTGCGATCTCGATAGAGGAGGCGCTGTCAGACGAGATCACAACCTCTAGAAAA 449
Qy  111 GluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrGlyThrPro 130
Db  450 GAGACTGATACTTGGCGGATATCGTTCGAGTCTCTGAGGATGACGACGAGACACT 509
Qy  131 IleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsn 150
Db  510 ATCCATGTGTTTGTACCAACACAGATCAGAGAGGACTTGATTCAGTGAATGTCGGTT 569
Qy  151 ValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaVal 170
Db  570 GCCTATAGACCATCGCATGCTGATGCAACTATGATCATGAGATGATGGTGTGAGATCAGTG 629
Qy  171 GlnGlyGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeu 190
Db  630 CAGGTGTGAGGAGAGATCTTCAGCTAGAGACCATTTGGAAGAGTTGCTCTCGGAGCTTG 689
Qy  191 ProLysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysVal 210
Db  690 GCCAAGAAAAATTTTGAAGCAATTTTCAGGAACTGAGATTTCTTGCCCTATGTCTCGCAAGTT 749
```

```
211 HisGlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIle 230
Db  750 CACCATGTTGTACTTCCAGAAAGAAATTTGGTAGACGACGAGAAATTTAACCTCTGGAACAGATA 809
Qy  231 GluSerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIle 250
Db  810 GAAAAATAACATTGTCAGATGCCCTAATCCCGAGTATCGGAAAGATGATAGCTGCGATT 869
Qy  251 AspArgValArgValArgGlyAspSerValGlyGlyValIleThrCysValAlaArgAsn 270
Db  870 GATGCTGTCTAGGACAAAGGGAACCTCTGTGTGGTGTGTGTGACCTGATTTGTCGGAAT 929
Qy  271 ValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAla 290
Db  930 GCTCCAGTGGGCTTGTGTACACCGGTTTTGATAAATCTTGAAGCAGAACTGGCAAAAGCT 989
Qy  291 MetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAsp 310
Db  990 TGTATGTCGCTACTCTGCAACTAAGGATTTGAGTTTGAAGCGGCTTTGCAGGTACCTTT 1049
Qy  311 LeuThrGlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThr 330
Db  1050 TTGACTGGTCTTGAACACAATGATGATGTTCTATACCGATGAAATGGAAGAATACGTACC 1109
Qy  331 ArgThrAsnArgSerGlyGlyValGlnGlyLysSerAsnValGluIleValHisPhe 350
Db  1110 AGAACCAACCGACTGTTGGTAATTCAGGAGGATCTCAATGTTGAATTAATAAATG 1169
Qy  351 LysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGlu 370
Db  1170 AGAGTAGCCTTCAAGCCAAACATCAACAATTTGGAAGGAAGCAAAATACGGTAACAGAGAC 1229
Qy  371 ArgGlnAsnValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaProArgAla 390
Db  1230 AAGGTAGAAAACCGAAATGATTTGCGCGTGGTGGTGCATGATCTTGTGTGTTCACGAGCT 1289
Qy  391 ValProValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisVal 410
Db  1290 GTGCCAATGTTGGAAGCAATGTTGGCTCTAGTCTTGTGATCAATTTGATGGCGCAATAC 1349
Qy  411 AlaGlnCysGluMetPheAlaLeuAsnThrAlaLeuGlnGluProVal 426
Db  1350 GCACAATGCCATTGTTTTCCAATAAATCCAGAGTTGCAGGAACCTCTC 1397

RESULT 10
AR432945 LOCUS
DEFINITION Sequence 11 from patent US 6653531.
ACCESSION AR432945
VERSION AR432945.1 GI:40195586
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 966)
AUTHORS Cahoon, R.E. and Falco, S.C.
TITLE Chorismate synthase from plants
JOURNAL Patent: US 6653531-A 11 25-NOV-2003;
FEATURES Location/Qualifiers
    source
        1..966
        /organism="unknown"
        /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.:      1.42e-75      Length:      966
Score:          1257.00      Matches:    245
Percent Similarity: 89.30%      Conservative: 22
Best Local Similarity: 81.94%      Mismatches: 26
Query Match:      57.14%      Indels:     6
DB:               6          Gaps:         2
```



```
QY      97 ----- 97
Db      83229 GTCTCTGAAAAATCCATTATTTCTGTCTGCACAGCTGGCTTCTGTCAAAACAGGATTA 83170
QY      97 ----- 97
Db      83169 TGCANAATGCAATCCTGCTAACCTCTTGCGAGAAATATGATTGTATGTTACATTACACCATTA 83110
QY      98 -----
Db      83109 ATGTTAACTGACATAAAATTTTGAATTATCATCAGACGGCCAGCCAGACAGATAAACACAC 83050
QY      107 rThrArgLysGluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyVal----- 125
Db      83049 CCCAAGAAAGGAGCTGACACTTGCANAAATTTCTTTTCAGGGACACATGAAGGTAT-GGTGT 82991
QY      125 ----- 125
Db      82990 CTTTCAGGACATGGGATAAACTTGGAAAAATATTTTTGAATTTAGAGTATAGCTTGTCTC 82931
QY      126 -----
Db      82930 TCATATTCACGATAGCATTTGATGGGTTTCTGCAGGAATGACCACTGGGACACCAATTCA 82871
QY      132 uValIleValProAsnThrAspGlnIleGly----- 142
Db      82870 TGTGTTTGTCCGGAACACAGATCAGAGAGGGGGGTGAAGTACCTAAGTTTTTTTCGCATTA 82811
QY      142 ----- 142
Db      82810 ATGTATTGAATACTACCTCCGTCCTCCAAAATATAACAACCTTTGGCTAGCAATCTGGACA 82751
QY      142 ----- 142
Db      82750 CACAATTTGCCAGATTCATAGCTAATAAATGCTTATATTTTGGGACAGAGGGAGTAACATA 82691
QY      143 -----
Db      82690 TTAAGCTGGAAATAACCATATTCATCCACAGGATTCACAGTGAATGGCTAAGGCCTA 82631
QY      152 rArgProSerHisAlaAspAlaThrTyraSpPheLysTyTGlyValArgAlaValGln-- 171
Db      82630 CAGACCTTCACATGCAGATGCNACTTATGACTTCAATACGGTGTATAGCAGTGCAGGT 82571
QY      171 ----- 171
Db      82570 ACTTTTTTTCATGTTTTTATATGTTGTAATATAGCCATTCATTTGAACCATCAACTGTAGTA 82511
QY      171 ----- 171
Db      82510 CCTAAACTTCTCTTATTTGAATGTCTAAACCTGCGGTACCTAAACCTTCTCTTACTTGAAT 82451
QY      172 -----
Db      82450 GCCTAACCAACATCTTCTATCAACCGAATGATAGGAGGTGGAGATCATCAGCAAGAGA 82391
QY      180 sThrValGlyArgValAlaAlaGlyValLeuProLysLysIleLeuLysCysG1 200
Db      82390 GACCATTGGAGGGTGGCTGTCAGGAGCTCTTGCAAGAAGAAATTTCTAAGCTCAATCTGG 82331
QY      200 Y----- 200
Db      82330 AGTAGAGGTGTGATCCTACTTGGCAACCTTTTACTCTTAACTGATCAAGATTTGTATCAA 82271
QY      200 ----- 200
Db      82270 GCATCGCAATTGACATGTTTACTTATCCATGACAACTTTGGACGCTGAAGAAAAATGAATC 82211
QY      200 ----- 200
Db      82210 AGCATTCGAATATACATCCTAAATATATGCTGAAGCCATGCACAAATTTCTGTCCAATGTCT 82151
QY      201 -----
Db      82150 CATAGCAGCAAAATATTTTCAAATTCAGATCTTGGCGTTTGTGTCCAAGGTGCATCAAGT 82091
QY      213 lValLeuProGluAspAlaValAspTyrGlySerValThrLeuGlu----- 228
Db      82090 TGTACTACCAAGAGATGCCGTTGATTATGACACTGTGAACATGGAACAGGTCTCTGGTTC 82031
QY      228 ----- 228
Db      82030 CCACCATTCTTCAAGTTTAAACACTTCTGGACAGCTCTCTGAAGGCATTTTGTCAATATTATA 81971
QY      229 -----
Db      81970 TTTATTTTATTTTCCAGATAGAAGCAACATTTGTAGATGTCCTGATCCCAAAATATGCAC 81911
QY      244 lLysMetIleAspAlaIleAspArgValArgValArgGlyAspSerValGlyValI----- 264
Db      81910 AGAAGATGATTGATGCAATCGAATAAGTACGAGTTAGAGGTGATTCCGATTGGTGTGTGG 81851
QY      264 leThrCysValAlaArgAsnValProArg----- 273
Db      81850 TCACATGCATTTGCAAGAAATGTTCTCGTGAAGTATTTAAACCTGCAGTTATGTATTAT 81791
QY      273 ----- 273
Db      81790 TTTGATATATGTCCTCCAGCTAAATAAATAAATAAAGTCATCTGTTTGGCACTCTCTGCT 81731
QY      273 ----- 273
Db      81730 TATGACCGCCTGCTTTTGGAAATTTGCACAAAATCTGTTTGAACATTGAAGCTTGATCAT 81671
QY      273 ----- 273
Db      81670 CATTAGAAAAATATTTTTTCTCTGCTGAGCTACTGATATGATAGACAGCAACAATTTGT 81611
QY      274 --GlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMetLeuS 293
Db      81610 AGGGGATGGCTCTCTGTTATTTGACAACTTGAGGCTGAATTTGGCGAAAGCTATGCTTT 81551
QY      293 erIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAla----- 307
Db      81550 CTCTCTCTGCAAGCAAGGGGTTTGAGATCGCAGTGCAGTGAATTTGACAGGTAATTTGAATTAATC 81491
QY      307 ----- 307
Db      81490 TGCACATAATTAGTTGATTATTTTTCGAAATAGCTGTGCACTGCCAACTACTTTGCTAC 81431
QY      308 -----
Db      81430 TGGCCTCAACTACTAATACTAATACTTGTGCCAGGTACTGACTACACTGGAAGTGAGGATA 81371
QY      317 snAspGluPheTyrMetAspLysAlaGlySerValArgThrArgThrAsnArgSerGlyG 337
Db      81370 ATGATGATGTTCTATATGATGAGGCTGGAATGTGAGAACAAAGAACTAAATCGTTCCAGCG 81311
QY      337 lyVal----- 338
Db      81310 GTGT-GCAGGTTTTTATCTAGTCCCTGTTGTTTACAGTTATATTATTTTATTGTTTCG 81252
QY      338 ----- 338
Db      81251 CTACATCTAGTGCATTTCTAGATTAAAAAATACTAGTAATAAGTATATAACACCGTTGTCT 81192
QY      339 -----
Db      81191 TAGTATTTTATAATGTGTATGTCCATGCAGGAGGGATCAAAATGGTGAATTTATATACT 81132
QY      350 heLysValAlaPheLysProThrProSerIleGlyVal----- 362
Db      81131 TCAAAGTAGCTTTTCAAGCCCAACCGCGACTATTGGGGT-GAGATATCCGGTTCAGTAAC 81073
QY      362 ----- 362
```

```
Db 81072 TCATTTTTTCATCAGTATATTAAAAAATCTATCTTAGAAATAGTCTATGTCATCT 81013
Qy 362 ----- 362
Db 81012 TTTTAAATGGCAAGCACATATGGTGTGTAGTCTTGTGGATGACGAGCTAAACCAAGTGT 80953
Qy 362 ----- 362
Db 80952 TCTCATAGTACATTTTGTGTCTTTTACGGTGTATCATATATTCATATGTAACTTTGAGG 80893
Qy 362 ----- 362
Db 80892 ATTTCATTCTTAAAGGCTATATTGCTACTTATTGTGATATTGGATTTCGCTTTTCC 80833
Qy 363 -----LysGlnAsnThrValSerArgGluArgGlnAsnValGluLeuLeuAlaArgGlyA 381
Db 80832 AGAAGAAACAACACTACTGTGTTCAAGGGAGCATGAGGATGTTGAACCTTTTAGCAAGGGCC 80773
Qy 381 IGHISAPPProCysValAlaProArg----- 389
Db 80772 GCCAGACCCATGTGTGTCCTCGCGGTAATGGATCTACGGCTCTACTACCTCCGCTCT 80713
Qy 389 ----- 389
Db 80712 AGAATAAGTTAATATAGTAGTACGGGATGTGATATATCCAGTATAACGAATTTGGTCAACAC 80653
Qy 389 ----- 389
Db 80652 CTATCCAGATTCGTAGTAGGATATGTTACATCCCGTACTAGATTAACTTATTCTGGAAG 80593
Qy 389 ----- 389
Db 80592 GAGGAGTACCTAAATTTTCTTCCCTTAACCACTCACTCATGATAGTATATCTGACAT 80533
Qy 390 -----AlaValProValValGluSerMetAlaAlaLeuVal 401
Db 80532 TTTCTCAATCAATTAACCTCTGTAGTGTGCGGATGGTGTGAGTCCATGCGCATTAGTC 80473
Qy 402 LeuMetAspGlnLeuMetAlaHisValAlaGlnCysGluMetPheAlaLeuAsnThrAla 421
Db 80472 CTCTGGACCACTGATGCGGCACATGCTCAATGTGAGATGTTTCCACTGAACCTTGCT 80413
Qy 422 LeuGlnGluProValGlySer 428
Db 80412 CTACAGAACCAAGTTGGCTCT 80392
```

RESULT 12
AP006487.1/c
WPCOMMENT

Sequence split into 6 fragments LOCUS AP006487 Accession AP006487

Fragment Name	Begin	End
AP006487_0	1	110000
AP006487_1	100001	210000
AP006487_2	200001	310000
AP006487_3	300001	410000
AP006487_4	400001	510000
AP006487_5	500001	523414

Continuation (2 of 6) of AP006487 from base 100001 (AP006487 Cyanidioschyzon merolae str

Alignment Scores:

Pred. No.:	8,75e-69	Length:	110000
Score:	1195.00	Matches:	244
Percent Similarity:	69.57%	Conservative:	60
Best Local Similarity:	55.84%	Mismatches:	114
Query Match:	54.32%	Indels:	19
DB:	2	Gaps:	5

US-10-677-179-8 (1-429) x AP006487_1 (1-110000)

```
Qy 3 ThrValPro-LysProGlnValAlaHisSerArgAlaArgLeuAlaProArgAla11 22
Db 84108 ACAGTACCGACGCCCC-----GCTGCTTCCATGTTGCTTCTTCCCTCTTGGCTG 84058
```

```
Qy 22 eGlyAlaLeuLeuGluPheAlaProAlaSerSerSerLeuArgPheAlaValHisArg-- 41
Db 84057 CCCTGGGCGGCTGCGAGCGGCGACAGACACTACCGAGTCTCTCTCCACAGAAA 83998
Qy 42 -----CysArgThrAlaArgLeuGluValLysAlaSe 52
Db 83997 TGGAAAGTCGCTCCGGGGAAGCTGTTTGC-----GCTCGTAGTGTGTTTCATGGTCT 83944
Qy 52 rGlyAsnThrPheGlyAsnThrPheGlnValAlaThrTyrGlyGluSerHisGlyGlyGly 72
Db 83943 CGGCAACTCTCTTTGGGAGGATATTTCGAATAACGACTTTTGGTGAATCCCAAGGTTGG 83884
Qy 72 yValGlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAspLeuGly 92
Db 83883 AGTCGGAGTTGTCTATGATGTTGTCACACCGCTGCACCTACTACTATAGAGAGCTTCA 83824
Qy 92 nValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLysGluTh 112
Db 83823 GATTGAGTTAAACAGCGGTAGACGGGACAGTCTAGGCTCAGCACACCGAACGAAAGC 83764
Qy 112 rAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThrProIle 132
Db 83763 AGATGAACCGGAGATCTCTCTGCGCTGCTCGACGGTCTCACTATGGGTACACCTATAGC 83704
Qy 132 uValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnValTy 152
Db 83703 TATCGTTGTTTGAACACACAGACCAAGGCTCTCAGGATATGAGGCAATGAAGGTAAAGCTA 83644
Qy 152 rArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGlnGly 172
Db 83643 TCGACCATCGCACGCGGATGCAACTATGATGCAAAAGTACGGCATACGCGCTCGCTGG 83584
Qy 172 yGlyGlyArgSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuProly 192
Db 83583 CGGTGGAGTTCATCAGCTCGGGAACAATAGGAGCTGTGCTGCGAGCGCATTCGCCAA 83524
Qy 192 sLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysValHisGly 212
Db 83523 AAAGATCTCTCGACATATGCGGGAACAACAATCTTCCTACGCTGGAAGGCGTGAAGCA 83464
Qy 212 nValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGluSe 232
Db 83463 TATCGAGCGCAGCAAT-----ATTGATCCCAACGCTGTTTCGTACGAGCAAAATCGAATC 83410
Qy 232 rAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIleAspAr 252
Db 83409 GAACATCGTTCGCTGCTGATTCAGAGCGAGCATCTCAAAATGATCCAGCTTATCGATGA 83350
Qy 252 gValArgValArgGlyAspSerValGlyValIleThrCysValAlaIleArgAsnValPr 272
Db 83349 AATTCGCCGAACAGATAACAGTGTGCGGGTGTGTTGAATGCGTTGCAAGAACGTCCTCC 83290
Qy 272 oArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMetIle 292
Db 83289 TGCCGGGTAGATCACCAGTGTGTTGACAAGCTCGAAGCTGAAGTTCGCAAGAGCCCTCAT 83230
Qy 292 uSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeuTh 312
Db 83229 GTCGCTTCACGATCGAAAGGCTTTGAAGTAGGCTCGGGTTCGGGGGCACTTTGCTTAAC 83170
Qy 312 rGlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThrArgTh 332
Db 83169 GGGTGTAGACACACAGATCCCTCTTATATAGATAGCAATGGTAGGACGCGAGCGGAGGA 83110
Qy 332 rAsnArgSerGlyGlyValGlnGlyIleSerAsnValGluIleValHisPheLysVa 352
Db 83109 GAACGTTCCGGAGGTATTCAAGGAGGCATCTCGAACGGAGAGAACATTTGTTGTCGCT 83050
Qy 352 lAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGluArgGly 372
Db 83049 CGCCTTCAACCGACCTCTACGATATACGCGGAACAGAAATACAGTACGCGCTGCGGTGA 82990
```

```
Oy 372 nAsnValGluLeuLeuAlaArgGlyVArGHIsAppProCysValAlaProArgAlaValPr 392
Db 82989 GCAAGTAGCTCTTCGGCGAGGGGAGACAGATCCCTGTGTGCTTCCTGCTGTCC 82930
Oy 392 oValValGluSerMetAlaLaLeuValLeuMetAspGlnLeuMetAlaHisValAlaGl 412
Db 82929 GATGCTAGAGCCATGGTCCGCTCTGTTGTTGATCACTGCTTCAACAGATTGCCA 82870
Oy 412 nCysGluMetPheAlaLeuAsn---ThrAlaLeuGlnGluProValGly 427
Db 82869 GTGTTTCGCTGCTTCATTGATCCATCAACGCGCACCGAACCTCTGGGA 82821

RESULT 13
LOCUS AP003583/3 339650 bp DNA linear BCT 28-NOV-2001
DEFINITION Nostoc sp. PCC 7120 DNA, complete genome, section 3/19.
ACCESSION AP003583 BA000019
VERSION AP003583.1 GI:17129939
KEYWORDS
SOURCE Nostoc sp. PCC 7120
ORGANISM Nostoc sp. PCC 7120
Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
REFERENCE
1 Kaneko,T., Nakamura,Y., Wolk,C.P., Kuritz,T., Sasamoto,S.,
Watanabe,A., Iriiguchi,M., Ishikawa,A., Kawashima,K., Kimura,T.,
Kishida,Y., Kohaza,M., Matsumoto,M., Matsuno,A., Muraki,A.,
Nakazaki,N., Shimpoto,S., Sugimoto,M., Takazawa,M., Yamada,M.,
Yasuda,M. and Tabata,S.
Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120
DNA Res. 8 (5), 205-213 (2001)
PUBMED 11759840
REFERENCE 2 (bases 1 to 339650)
Kaneko,T.
Direct Submission
Submitted (02-MAY-2001) Takakazu Kaneko, Kazusa DNA Research
Institute, The First Laboratory for Plant Gene Research; Yana
1532-1, Kisarazu, Chiba 292-0812, Japan
(E-mail: kaneko@kazusa.or.jp,
URL: http://www.kazusa.or.jp/cyanobase/,
Tel: 81-438-52-3935 (ex.2338), Fax: 81-438-52-3934)

FEATURES
source
1. 339650
/organism="Nostoc sp. PCC 7120"
/mol_type="genomic DNA"
/db_xref="taxon:103690"
/note="synonym: Anabaena sp. PCC7120"
complement(98..2698)
/gene="al10594"
complement(98..2698)
/gene="al10594"
/note="ORF ID: al10594"
/codon_start=1
/transl_table=11
/product="aminopeptidase"
/protein_id="BAB72552.1"
/db_xref="GI:17129940"
/translation="MPHSYFDDQSDNAHKSFLPGARPHYPNDRPGQVEHIFLDNLND
IPQSYEGSCSIRLLPIRNGIDRLDRAVNIQIRSVQDVQDFDYDGEOLVQLSQ
PTEIGKRLIAIAYAVEKPGRIYFIQPSHYPNKVQVMTQGEDDSRFWPCFDYF
GOLSTSEIRVVPKSLMAISNGELIDTOEDGKHLYHWSQQIHPYTLMTLAVADPAE
IRDEWRGKPVYVYKSGREADAKRSKGTPTRMIEFLSKYGYAYAPFKYAOVSDFI
FGGMENTSTLLTDCLLDERAALDNRTESLVVHELAFQWGLDLVKGWSHAWIKE
GNASYSEVMTQEYSAEAAAYRLLEARSYLSYSDSGRYRPMVTHVYREAIEIYDRH
TYEGKSYHMIARELGBDLFWQAIQTFRDNAKHTVETIDLLRAIEKATGRNLAFLE
DOYVYRGHPDPKAYASWDGEAKLAKVTVTOVPADKQDVFLRIPIGFGNOTEA
AAKLTPTVVRVHERSQSFVPLTEKPDFISDVGNNHMLKTVSYLPIPELKAQASDP
DPSIRYAAALAKGGLKAYKALASLAKNEPFGVRYEVAKQALIKDDAOFGLVE
GLADKNVYRRAVQALAKIKTLASVKAVGVVKGODGASYESHSAACLAGAIRAINF
DDKSKEBAIKLLQSLBERAGNVEVSGAIGGLAEKLTSEAADLLILEYTKLGVPQ
FLRLAAIRALGKISFGQSPVNLERI IARLGELSKENFFILMAVVAALGQMETPKAMG
VLRSLADQTADGRVRFABEELANVQKNIGTIDNALRKLREBFDQLKQONQBLKRLN
```

```
LEAKSOTVKQ"
complement(2862..3059)
/gene="as10595"
complement(2862..3059)
/gene="as10595"
/notes="ORF ID: as10595"
similar to cytochrome d oxidase chain I"
/codon_start=1
/transl_table=11
/protein_id="BAB72553.1"
/db_xref="GI:17129941"
/translation="MRHCVHPILLSLPLYKIIILNIQWKSPDRPHMVGLIYYAFRVM
IAIGFFLAGLMVSLTFFEV"
complement(3151..3810)
/gene="al10596"
complement(3151..3810)
/notes="ORF ID: al10596"
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB72554.1"
/db_xref="GI:17129942"
/translation="MGVCHSKDVSSEFCLEGRPMFEFVIKDGYSKLGKLLLTADGECYVK
LSQLRPTFDWRLPTGTLQVGEKKYDAKTGKVTLKAEAVMAARTETSSAVKQVPM
ENTKPKPEKAKATILVCQKSDCKMRGKALCOALEATLSDRGLEDOVTIKGTGCMKC
KAGPNLVMPDKTRYTRIQAQOVFKLWNKHFDSSQSDHSHQAVISINTKFEPTFRVA
F"
complement(3877..4134)
/gene="as10597"
complement(3877..4134)
/gene="as10597"
/notes="ORF ID: as10597"
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB72555.1"
/db_xref="GI:17129943"
/translation="MKQFSSEVEKSVVEVDWADRMVOYORLOLEIPCWCEANQPLK
VELVTPMTVVQLMSVMRQFTSSRQDLIYSLEINWQSRYPIS"
4451..5146
/gene="alr0598"
4451..5146
/gene="alr0598"
/notes="ORF ID: alr0598"
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB72556.1"
/db_xref="GI:17129944"
/translation="MYRTMRPLLSTLSLVLLQGFVPIVOAQSGENTGDVASTAIQOV
IAANMTNSADGNFYPERLMNRALASILVKAFRLNQEANQPKSITVADVPRSHWA
YQDIQTVLKTDVMKRYGNLFFPNQVRTRAEALAIFAQYGVGFQSDVVNEILAPYP
DANSIPNWARRAIATVITEGINTDAQNNITPLRPMTRGDMAYLLSKYLQORQKQPD
PVVPTTDSQPLP"
5596..7503
/gene="alr0599"
5596..7503
/gene="alr0599"
/notes="ORF ID: alr0599"
/codon_start=1
/transl_table=11
/product="1-deoxy-xylulose 5-phosphate synthase"
/protein_id="BAB72557.1"
/db_xref="GI:17129945"
/translation="MHLSEITHPNQLHLSIRLQQLIARQIRDKHLOTVAEFGHLP
GLGVNFTLGLYOTDLDRDKVWDVGHQYAPHKLIICRYSDHFTLRQKQDIAGLYKR
GENKDFHFGAGHASTSI SAALGALARDMNGEKFKAVALIGDGAITGCMALAEANHAG
HLKPTLLVVLNQNENSI SPNVGAI PLYLNKMRSLSPVQRIKONFQFQKHIPVGES
LSBELGRIKGMKRLAVPKYGAPEELGFTYMGVDGHNLEELIATFQQAHQIAGPVL
VHVATIKGQYELAEKQDVGHQATPFNLITGKAIPSNKPKPPAIKAYFSGHVLKLAE
ONPKIIGITAMATGTGLQAKLPNQYIDVGAEGHAYVTLAAGLATGEGMRPVAAIY
STFLQRAYDQIITHDVCITQNLFPVFFCLDRAGIVGSDGPTHQWYDIAYLRCTIPNIVMA
```



```
Qy 331 ArgThrAsnArgSerGlyGlyValGlnGlyGlyIleSerAsnValGluIleValHisPhe 350
Db 231527 GTAACCTAACCGTTCGGTGGGATTCAGCGGTATTGCCAACGGAGAGAAATATCATTTTG 231468
Qy 351 LysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGlu 370
Db 231467 CGAGTGTGATTTAAACCCACGCCACGATTAGAAAAGACAAAACAGTTACTCGTGAA 231408
Qy 371 ArgGlnAsnValGluLeuAlaArgGlyArgHisAspProCysValAlaProArgAla 390
Db 231407 GCGGAAGAACCCATTATTAGCCGCAAGAGTAGACATGATCCGTGTCTTCTACCCCGTGCA 231348
Qy 391 ValProValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisVal 410
Db 231347 GTCCCGATGGTAGAAGCAATGTGCGTGTGCTACTGTCGATCATTTGTTAGCCCATCAT 231288
Qy 411 AlaGlnCysGluMet 415
Db 231287 GGACAGTGTAGGTG 231273

RESULT 14
SSRPAROC
LOCUS Synecocystis sp. rplL and aroC genes. linear BCT 05-JUL-1995
ACCESSION X67516
VERSION X67516.1 GI:452796
KEYWORDS chorismate synthase; ribosomal protein L12.
SOURCE Synecocystis sp. PCC 6803
ORGANISM Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
REFERENCE
AUTHORS Schmidt, J., Bubunenko, M. and Subramanian, A. R.
TITLE A novel operon organization involving the genes for chorismate synthase (aromatic biosynthesis pathway) and ribosomal GTPase centet proteins (L11, L1, L10, L12: rplKAJL) in cyanobacterium Synecocystis PCC 6803
JOURNAL J. Biol. Chem. 268 (36), 27447-27457 (1993)
MEDLINE 9408656
PUBMED 7505271
REFERENCE
AUTHORS Schmidt, J.
TITLE Direct Submission
JOURNAL Submitted (24-AUG-1992) J. Schmidt, MPI f Molekulare Genetik, Abt Wittmann, Imnestrasse 73, D-1000 Berlin 33, FRG
FEATURES
source
1..1884
/organism="Synecocystis sp. PCC 6803"
/mol_type="genomic DNA"
/strain="PCC 6803"
/db_xref="taxon:1148"
/clone="LCS IV"
/haplotype="haploid"
/clone_lib="EMBL3"
1..39
/gene="rplL"
1..39
/gene="rplL"
/codon_start=1
/transl_table=1
/product="ribosomal protein L12"
/protein_id="CAA47854.1"
/db_xref="GI:452797"
/db_xref="GOA:P23349"
/db_xref="UniProt/Swiss-Prot:P23349"
/translation="QIEBAGGKAAVK"
40..245
/misc_feature
246..1334
/gene="aroC"
246..1334
/gene="aroC"
/EC_number="4.6.1.4"
```

```
/codon_start=1
/transl_table=11
/product="chorismate synthase"
/protein_id="CAA47855.1"
/db_xref="GI:452798"
/db_xref="GOA:P23353"
/db_xref="UniProt/Swiss-Prot:P23353"
/transl_table="MGNTFGLFRIITRGESHGGVGVIIIDGCPRLLEISPERIQVDL
DRRPGSKITTPRKADQCEILISGVFEKTLGTPTALIVRNKDAKSDYNEMAYKIR
PSHADATYEAQYGIKRWQGGSSARETIGRVAAGAIAKILQAFNGVEIVAYVKSIQ
DIADATSVNTVLEQWNSVIRCPDECAEKMIERIDQVLRQKDSIGSVVECAIRNAP
KIGPEPVFDLAKMMSLPATKGFEGSGFAGTLLTGSQHNDEYIYDEAGEMWT
RTNRSGVQCGIISNGRPIIMRIAPKPTATIQEQKTVTSNIGBETTLAAKGRHDPVLP
RAYPMVEAMALVLCDELLRFPQACKTL"
1335..1378
/notes="spacer"
1379..1624
/notes="unnamed protein product; orf81"
/codon_start=1
/transl_table=11
/protein_id="CAA47856.1"
/db_xref="GI:452799"
/db_xref="UniProt/Swiss-Prot:P39204"
/transl_table="MLPQAIESSGQNCSSIARVWKIGRKFTNINRVNGLKLANPSKF
RCLMPRSPVGTMTPTKLTLLCPNSLPFCGDRARIQPI"

ORIGIN
Alignment Scores:
Pred. No.: 8,28e-71 Length: 1884
Score: 1191.50 Matches: 225
Percent Similarity: 77.90% Conservative: 57
Best Local Similarity: 62.15% Mismatches: 77
Query Match: 54.16% Indels: 3
DB: 1 Gaps: 1

US-10-677-179-8 (1-429) x SSRPAROC (1-1884)
Qy 53 GlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGlyGly 72
Db 249 GGAACACGTTTGGCAGTTATTTCGCATTACCACTTCGCGTAGTCCCATGGGGCGGC 308
Qy 73 ValGlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAspLeuGln 92
Db 309 GTGGGGTAAATTATTGATGTTGCTCCCTCCCGATTGGAAATTTCCCCAGAGAAATCAA 368
Qy 93 ValGluLeuAspArgArgArgProGlyGlnSerArgIleThrSerThrArgLysGluThr 112
Db 369 GTAGACCTAGACCGCGCGCGCCAGGACAGAGCAAAATATACCACTTCGTAAGAAAGCA 428
Qy 113 AspThrCysValIleLeuSerGlyThrHisGluGlyValThrThrGlyThrProIleLeu 132
Db 429 GATCAATGCGAAATTCCTCTGGCGTGTTCAGGGTAAACTCTGGGTACACCATGCC 488
Qy 133 ValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnValTyr 152
Db 489 ATTTTGTGCGAATAAGGATGCCCGTTCACAGATTACAATGAGATGCGGTGAATAT 548
Qy 153 ArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGlnGly 172
Db 549 CGTCTTCCACGCGCATGCCACCTATGAAGCGAAATATGTCATTTCGCAATGTCGAAGGA 608
Qy 173 GlyClyArgSerSerGlyArgLysThrValGlyValAlaAlaGlyAlaLeuProLys 192
Db 609 GGAGTCGCTCTTCTGTAGGAAACCATTTGCCGGGTGGCCGCGGCGCATCGCCAAA 668
Qy 193 LysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysValHisGln 212
Db 669 AAGATTCGGCCCAAGTTTAACGGGGTGAATTTGGCCTATGTGAAAGATTCAGAGAC 728
Qy 213 ValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGluSer 232
Db 729 ATT-----GAGGCGACGCGTGACAGCAATACGGGTGACCTTTGGAGCAAGTGGAAAGC 779
```


QY 233 AsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIleAspArg 252
 Db 780 AATATTGTTGTTGTTCCGATGAAGATGTCGCGAAATGATTGACCCATTGACCG 839
 QY 253 ValArgValArgGlyAspSerValGlyGlyValIleThrCysValAlaArgAsnValPro 272
 Db 840 GTCTACGACAAAGAACTCATTTGGCGGGTGTGGAATGTGCGATTCCTACGCTCC 899
 QY 273 ArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMetLeu 292
 Db 900 AAGGGTTGGGGAACACGATTGTGATGTTGAAGCGGATTTGGCAAGGCAATGATG 959
 QY 293 SerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeuThr 312
 Db 960 TCCCTCCCGCTACCAAGGGTTTGAAGTTGGCTCTGTTTGGCGGTACTTTACTGACT 1019
 QY 313 GlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThrArgThr 332
 Db 1020 GSTAGTCAACATAATGATGATGACTACTCTGATGAAGCGGGGAATGGCGTACTCGCAC 1079
 QY 333 AsnArgSerGlyGlyValGlnGlyGlyIleSerAsnValGluIleValHisPheLysVal 352
 Db 1080 AACCGTTCTGGGGGGTTCAGGAGGATCAGCAACGGCAACCCATTATTATGCGCAT 1139
 QY 353 AlaPheLysProThrProSerIleGlyValLysGluAsnThrValSerArgGluArgGln 372
 Db 1140 GCATTTAAACCTACTGCCACCATTTGTCAGACCAAAACCCGTTAGTAACATGGGAG 1199
 QY 373 AsnValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaProArgAlaValPro 392
 Db 1200 GAAACCACTTACCGCTAAAGGTGCGCATGATCTCTGCTTTGCGCGGGCTGTGCC 1259
 QY 393 ValValGluSerMetAlaAlaLeuValLeuMetAspGluLeuMetAlaHisValAlaGln 412
 Db 1260 ATGTGGAGCGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1319
 QY 413 CysGlu 414
 Db 1320 TGTAAA 1325
 RESULT 15
 D90906/c
 LOCUS
 DEFINITION Synchocystis sp. PCC 6803 DNA, complete genome, section: 8/27,
 920916-1056466
 ACCESSION D90906 AB001339 BA000022
 VERSION D90906.1 GI:1652492
 KEYWORDS
 SOURCE Synchocystis sp. PCC 6803
 ORGANISM Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
 REFERENCE
 AUTHORS Kaneko, T., Tanaka, A., Sato, S., Kotani, H., Suzuki, T., Miyajima, N.,
 Sugita, M., and Tabata, S.
 TITLE Sequence analysis of the genome of the unicellular cyanobacterium
 Synchocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
 region from map positions 64% to 92% of the genome
 JOURNAL DNA Res. 2 (4), 153-166 (1995)
 MEDLINE 96127529
 PUBMED 8590279
 REFERENCE
 AUTHORS Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y.,
 Miyajima, N., Hirose, M., Sugita, M., Sugiura, M., Sasamoto, S., Kimura, T.,
 Hosouchi, T., Matsuoka, A., Muraki, A., Nakazaki, N., Naruo, K.,
 Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A.,
 Yamada, M., Yasuda, M., and Tabata, S.
 TITLE Sequence analysis of the genome of the unicellular cyanobacterium
 Synchocystis sp. strain PCC6803. II. Sequence determination of the
 entire genome and assignment of potential protein-coding regions
 JOURNAL DNA Res. 3 (3), 109-136 (1996)
 MEDLINE 97061201
 PUBMED 8905231
 REFERENCE
 AUTHORS 3 (bases 1 to 135551)

AUTHORS TITLE JOURNAL

Tabata, S.
 Direct Submission
 Submitted (28-JUN-1996) Satoshi Tabata, Kazusa DNA Research
 Institute, The First Laboratory for Plant Gene Research; Yana
 1532-3, Kisarazu, Chiba 292-0812, Japan
 (E-mail: tabata@kazusa.or.jp, URL: http://www.kazusa.or.jp/cyano/,
 Tel: 81-438-52-3933 (ex.2330), Fax: 81-438-52-3934)
 Potential protein coding regions were assigned on the basis of
 similarity search of the ORFs and GeneMark analysis.

COMMENT FEATURES source

Location/Qualifiers
 1. 135551
 /organism="Synchocystis sp. PCC 6803"
 /mol_type="genomic DNA"
 /strain="PCC6803"
 /db_xref="taxon:1148"
 /note="synonym: Synchocystis PCC6803"
 /complement (27. .281)
 /gene="ssl3712"
 /complement (27. .281)
 /note="orf ID: ssl3712"
 /unknown protein
 /codon_start=1
 /transl_table=11
 /protein_id="BAA17414.1"
 /db_xref="GI:1652493"
 /translation="MGETKLADQRWQTHDCKDYLSYLHRRYELVEVAVSFWP
 HVPPEKTLTLLAWEXELILKWRSPFNKECWMMWQPFGLKLS"
 /complement (634. .1722)
 /gene="aroc"
 /complement (634. .1722)
 /note="orf ID: ssl1747"
 /codon_start=1
 /transl_table=11
 /product="choiximate synthase"
 /protein_id="BAA17415.1"
 /db_xref="GI:1652494"
 /translation="MGNTFGSLFRITTFBHGSGGVVIGDCPPRLRISPEIQLD
 DRRPGSKITTPRKADQCILSGVFEGKLTGPIALVLRNKDARSQDYNEMAYKR
 PSHADATYAAKIRGNWGGRRSARETIGRAAGAKILAQPNQVEIVAYKSIQ
 DIATVNSNTVTLEQVSNIVRCDECAERIERIDQVLRKQDSIGVVEICARNA
 KGLGEPVDFKLEADLAKAMSLPATKPFEGSGAGTLLTGSHNDYLLDEAGSWRT
 RTNSGGVGGISNGEPLIMEIAKPTATIGQEQKTSVNIQEETTLAKGRHDPCLP
 RAVPMVEAMALVLCDDLRLRQAQCKTL"
 /complement (1929. .2315)
 /gene="rpl12"
 /complement (1929. .2315)
 /note="orf ID: ssl1746"
 /codon_start=1
 /transl_table=11
 /product="50S ribosomal protein L12"
 /protein_id="BAA17416.1"
 /db_xref="GI:1652495"
 /translation="MSRATDQILEQLSLLEASLVKQIEEAFGVSAAAPVGMW
 AAAAFAAEAEKTEFDFVLEVEFPADKIAVLKVRITITGLKAEKELVESTPKAI
 KEATKDDAEAIKKQIEAGKAAVK"
 /complement (2403. .2924)
 /gene="rpl10"
 /complement (2403. .2924)
 /note="orf ID: ssl1745"
 /codon_start=1
 /transl_table=11
 /product="50S ribosomal protein L10"
 /protein_id="BAA17417.1"
 /db_xref="GI:1652496"
 /translation="MGRTRENKATVSDVQLFQDAQMTVIDYQGLTVAEITDLNR
 LRLGGTCKIAKNTLVRLRAGQAWSPMEFLTGTAILVLKEDLGAIKAYKKFK
 DTKTELRGGVLEKSLTQADVEAIGDLPSEQLMGQTAGGINALATKIALGIKEVA
 SVARGLOAHVDKE"
 /complement (3171. .3887)
 /gene

CDS	/gene="rpl11" complement(3171..3887) /gene="rpl11" /note="ORF_ID:s111744" /codon_start=1 /transl_table=11 /product="50S ribosomal protein L1" /protein_id="BAA17418.1" /db_xref="GI:1652497" /translation="MTKKLSKRMQAATAKVDSDSKLSPLEAMELLKETATAKDETAE AHRLGIDPKYSQDQIRTVSLPGTGTQVRVAVLARGEKVKENTADAGADIAGEBELI SEIQKQMDFDVLATPMPMKIARLQGLGPRGMPSPKGTVTADLAAAVNEFKAG KLEFRADRTGTVHMFGRKASPSADLLANLKAQETIDRNRPAGKGRFWRITVPSSS MGPSIPVDINALRDLKPEDN" complement(3974..4399) /gene="rpl11" complement(3974..4399) /gene="rpl11" /note="ORF_ID:s111743" /codon_start=1 /transl_table=11 /product="50S ribosomal protein L11" /protein_id="BAA17419.1" /db_xref="GI:1652498" /translation="MAKKVVALIKALPAGKANPAPPVPGALGOHGVNINAFCKEYNA KQADKPMIIEVESVIFEDRSFTFLKTPPASVLIRKAAAGVKGSSSPNKNKVASITR BOLREIAOTKLPLDNLNANDIDAAMNIEGTARNMGITVNS" complement(4501..5118) /gene="nusG" complement(4501..5118) /gene="nusG" /note="ORF_ID:s111742" /codon_start=1 /transl_table=11 /product="transcription antitermination protein; NusG" /protein_id="BAA17420.1" /db_xref="GI:1652499" /translation="MSFTDDQSPVAEQNKTPPSBGHWFAVOVASGCEKRVKLNLEQRI HTLDVADRILOVEIPKPTPIVKIRKDQVQGEKIPFYVLIRIMDDDAQVQVKNPT HVINFVSGEQKHGRGRGHVLPMLSHGEVERIFRHVDEGEPVVKIDMEIGDHIMVL SGPKFDGSDVIEVSPERSKULALISIPGRETPVELLEFTQVEKQ" complement(5398..5643) /gene="secE" complement(5398..5643) /gene="secE" /note="ORF_ID:s13335" /codon_start=1 /transl_table=11 /product="secretory protein; SecE" /protein_id="BAA17421.1" /db_xref="GI:1652500" /translation="MWKKEAVRTSDSTENSDVNQARSNFTAAATKDELAKVWPSRQ LISESVAILMVLVSTVIFYVDQIFGWITTKQPLFG" complement(5755..5827) /note="trnW-CCA" complement(5945..6313) /gene="rpl19" complement(5945..6313) /gene="rpl19" /note="ORF_ID:s111740" /codon_start=1 /transl_table=11 /product="50S ribosomal protein L19" /protein_id="BAA17422.1" /db_xref="GI:1652501" /translation="MTMNAQAIIINSIEAFLEKEDLPTTHVGDTIKVGKIVGEGKERI QPYETGVIAKXNGGISETITVYRKIPQGVGVVERVLLHSPRVASIKVLRRGKVRRAKLY VLRDVGATRIKQRFDRAL" 6418..6843 /gene="slr1826" 6418..6843 /gene="slr1826" /note="ORF_ID:s1r1826"	gene	unknown protein" /codon_start=1 /transl_table=11 /protein_id="BAA17423.1" /db_xref="GI:1652502" /translation="MQFCWIIIEKPMTTITSOAIARYRDQLAHCPENMAQDITIEDCE GNLEDAALTLGIQVGQPPDNDMLGLEAKRCRAICGEGPPAALKDNINISTVVNHLSE RKVCHPLLAFLVAVYLEQGIIDDFCAPIFHKLNDGNN" 7044..7835 /gene="slr1827" 7044..7835 /gene="slr1827" /note="ORF_ID:s1r1827" unknown protein" /codon_start=1 /transl_table=11 /protein_id="BAA17424.1" /db_xref="GI:1652503" /translation="MSFFAQHTSVVYATIDLLFGGSKRKADEYRSAYSWADOVRDFCREI IGCPTVLVGNISIGSLVALTAAESPFLVAGLIMASLPDVSRLQQAAPFLRPWIEKVE SAFSPWLLNGLFKFVRPPIIKRWAGLATGLRQEQDLDFEELVQIICSPQDGLG AEVAFQOLFSSVRQSHFAPSVKELLPTLDLPIILLMGDGRFIPCRYQQGFALLNKKI EFQLWPGVGCHLHDECPNFKNQTCITLWLRNFDKSFELSEPVVTA"	gene	gene CDS	Alignment Scores: Pred. No.: 1.96e-68 Length: 135551 Score: 1191.50 Matches: 225 Percent Similarity: 77.90% Conservative: 57 Best Local Similarity: 62.15% Mismatches: 77 Query Match: 54.16% Indels: 3 DB: 1 Gaps: 1 US-10-677-179-8 (1-429) x D90906 (1-135551)	Qy	53 GlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGlyGlyGly 72	Db	1719 GGAAACACGTTTGGCAGTATTATTCGACCTTACCACCTTCGGTGAGTCCCATGGGGCGGC 1660	Qy	73 ValGlyCysValIleSerGlyCysProProArgIleProLeuThrGluAlaAspLeuGln 92	Db	1659 GTGGGGGTAAATTATTTAGTGTGCCCTCCCGATTGGAAATTTCCCCAGAGGAATCCAA 1600	Qy	93 ValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLysGluThr 112	Db	1599 GTAGACCTAGACCGCGCGCGGCGGACAGACAGCAAAATACCACCTCGTAAGGAAGCA 1540	Qy	113 AspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThrProIleLeu 132	Db	1539 GATCAATGCGAAATTCCTCTGGCGGTTCGAGGGTAAAACTCTGGGTACACCGATCGCC 1480	Qy	133 ValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnValTyr 152	Db	1479 ATTTTGTGGCAATATAGGATGCCCGTTCCAGGATTACAAATGAGATGGCGGTGAAATAT 1420	Qy	153 ArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGlnGly 172	Db	1419 CGTCTCTCCACGCCGCGATGCCACCTATGAACGGAATATGGCATTCGCATTTGCAAGGA 1360	Qy	173 GlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuProLys 192	Db	1359 GGAGGTGCTCTTCTGTAGGAAACCAATTTGGCGGGTGGCGCGGCGGCGATCGCCAAA 1300	Qy	193 LysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysValHisGln 212	Db	1299 AAGATTCTGGCCCCAGTTTAAACGGGGTGGAAATTTGGCCCTATGTGAAAGTATCCAGAC 1240	Qy	213 ValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGluSer 232	Db	1239 ATT-----GAGGCGCGGTGGACAGCAATACCGTGACCTTGGACCAAGTGGAAAGC 1189	Qy	233 AsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIleAspArg 252	Db	1188 AATATTGTTCTGTTCCCGATGAAGAATGTGCCCGAAAAAATGATTGAACGATTTGACCAG 1129
-----	--	------	--	------	-------------	--	----	--	----	--	----	--	----	--	----	--	----	---	----	--	----	--	----	--	----	---	----	--	----	---	----	--	----	---	----	--	----	--	----	--	----	---	----	--	----	--

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 27, 2005, 01:03:39 ; Search time 619 seconds

(without alignments)
4102.696 Million cell updates/sec

Title: US-10-677-179-8

Perfect score: 2200

Sequence: 1 MTTVPKQQVAHSRLAPR.....VAQCEMFALNTALQBPVGSF 429

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of Hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlp
-Q/cn2_1/USFTO.spool_p/US10677179/runat_25082005_150239_24825/app_query.fasta_1.583
-DB=N Geneseq 16Dec04 -QFT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOFC=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLCN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10677179 @CN_1_1_708 @runat_25082005_150239_24825 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 16Dec04:
1: Geneseqn1980s:
2: Geneseqn1990s:
3: Geneseqn2000s:
4: Geneseqn2001as:
5: Geneseqn2001bs:
6: Geneseqn2002as:
7: Geneseqn2002bs:
8: Geneseqn2003as:
9: Geneseqn2003bs:
10: Geneseqn2003cs:
11: Geneseqn2003ds:
12: Geneseqn2004as:
13: Geneseqn2004bs:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2200	100.0	1626	3	Aaz50438 Corn clon
2	1797.5	81.7	1835	3	Aaz50435 Corn clon
3	1461	66.4	1595	3	Aac40227 Arabidops
4	1454	66.1	1311	12	Adn73862 Thale cre
5	1454	66.1	1311	13	Adt91426 Arabidops

6	1257	57.1	966	3	Aaz50440	Aaz50440 Rice clon
7	1191.5	54.2	1089	13	ADS47934	Ads47934 Bacterial
8	1186.5	53.9	1086	13	ADT44117	Adt44117 Bacterial
9	1182.5	53.8	1086	13	ADT45459	Adt45459 Bacterial
10	1066	48.5	1015	3	AAZ50437	Aaz50437 Wheat cto
11	1035	47.0	1080	13	ADT46302	Adt46302 Bacterial
12	981	44.6	1068	13	ADS58432	Ads58432 Bacterial
13	938	42.6	1131	13	ADT47171	Adt47171 Bacterial
14	934	42.5	1137	12	ADP98700	Adp98700 C. albica
15	929.5	42.2	1239	13	ADR85489	Adr85489 Aspergill
16	888.5	40.4	1074	8	ACA24777	Act24777 Prokaryot
17	869.5	39.5	1421	13	ADS49229	Ads49229 Bacterial
18	862	39.2	1086	8	ACA54193	Acas54193 Prokaryot
19	860	39.1	1086	8	AAS52498	Aas52498 E. coli D
20	860	39.1	1086	8	ACA32569	Acas32569 Prokaryot
21	860	39.1	1086	13	ADT48732	Adt48732 Bacterial
22	860	39.1	3100	8	ACC79525	Acc79525 Aroc gene
23	859	39.0	1134	11	ACH99580	Ach99580 Klebsiell
24	858.5	39.0	1071	12	ADO57719	Ado57719 Actinobac
25	858	39.0	1056	8	ACA36973	Acas36973 Prokaryot
26	857	39.0	1086	4	AAS53974	Aas53974 Klebsiell
27	855	38.9	1083	8	ACA35743	Acas35743 Prokaryot
28	850	38.6	1083	8	ACA45011	Acas45011 Prokaryot
29	850	38.6	1089	10	ADR99834	Adr99834 Bacterial
30	850	38.6	1363	13	ADR84902	Adr84902 Aspergill
31	850	38.6	7363	13	ADR84315	Adr84315 Aspergill
32	847	38.5	1239	4	ADAO6952	Adao6952 aroc gene
33	847	38.5	1690	2	AAZ22883	Aaz22883 E. coli w
34	847	38.5	4530	4	AAO06958	Aao06958 pME4 comp
35	846.5	38.5	1074	8	ACA42924	Acas42924 Prokaryot
36	845.5	38.4	34980	13	ADT05648	Adt05648 Haemophil
37	845	38.4	110000	6	ABA92787_0	Abas92787 Buchnera
38	845	38.4	110000	6	ABA92787_1	Abas92787 Buchnera
39	843	38.3	1080	13	ADT46404	Adt46404 Bacterial
40	840.5	38.2	1071	5	AAF94408	Aaf94408 Haemophil
41	840.5	38.2	1071	6	ABK64972	Abk64972 DNA encod
42	840.5	38.2	1074	4	AAS53247	Aas53247 Haemophil
43	840.5	38.2	1074	8	ACA34045	Acas34045 Prokaryot
44	840.5	38.2	1077	13	ADT45862	Adt45862 Bacterial
45	840.5	38.2	110000	2	AAT42063_02	Continuation (3 of

ALIGNMENTS

RESULT 1
ID AAZ50438 standard; cDNA; 1626 BP.
AC AAZ50438;
DT 18-MAY-2000 (first entry)
DE Corn clone csiln.pk0050.d11 encoding chorismate synthase.
KW Corn cDNA clone csiln.pk0050.d11; chorismate synthase;
KW immunological screening; herbicide resistance; antibody; gene mapping;
KW corn; ss.
OS Zea mays.
FH Key Location/Qualifiers
CDS 79..1368
FT /*tag= a
FT /product= "Chorismate synthase"
PN WO200005353-A2.
XX PD 03-FEB-2000.
XX PF 20-JUL-1999; 99WO-US016353.
XX PR 21-JUL-1998; 98US-0093611P.
XX

PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX Cahoon RE, Falco SC, Pember SO;
XX WPI; 2000-182687/16.
DR P-PSDB; AAY44893.
XX
XX New chorismate synthase polypeptides used to alter the level of the
PT enzyme and thus the level of aromatic to non aromatic amino acids in
PT transformed plants.
XX
XX Claim 3; Page 32-33; 39pp; English.
XX
XX The present sequence is corn cDNA clone designated csiln.pk0050.d11
CC encoding chorismate synthase. This clone is isolated from csiln cDNA
CC library which is derived from corn silk tissue seedling. Chimeric gene
CC comprising this cDNA operably linked to regulatory sequences is used to
CC transform host cells to alter the level of expression of chorismate
CC synthase. The gene and its products may be used for immunological
CC screening of cDNA expression libraries and to create transgenic plants
CC which may also be herbicide resistant. Synthetic peptides derived from
CC the gene are to raise antibodies, and used in screening assays to
CC identify inhibitors which may be useful as herbicides. Probes and primers
CC from chorismate synthase gene are used for gene mapping and as markers
CC for traits linked to those genes
XX
SQ Sequence 1626 BP; 405 A; 382 C; 439 G; 400 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 7,986-193 Length: 1626
Score: 2200.00 Matches: 429
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-677-179-8 (1-429) x AA250438 (1-1626)

QY 1 MetThrValProLysProGlnValAlaHisSerArgAlaArgLeuAlaProArg 20
DB 79 ATGACGACCGTCCCAAGCCACAGAGGTGGCGCACTCACGGGACGGCTCGCACCCGCG 138
QY 21 AlaIleGlyAlaLeuLeuGluPheAlaProAlaSerSerSerLeuArgPheAlaValHis 40
DB 139 GCGATCGGCGCTTGTGAGTTTGGCCCGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 198
QY 41 ArgCysArgThrAlaArgLeuGluValLysAlaSerGlyAsnThrPheGlyAsnTyrPhe 60
DB 199 CGCTGCGGCACTGCTCGCTAGAGGTGAAGGCATCTGGAAACACCGTTTGGAAACTACTTT 258
QY 61 GlnValAlaThrTyrGlyGluSerHisGlyGlyValGlyCysValIleSerGlyCys 80
DB 259 CAGGTTCGAACCTATGTTGTAATCTCATGGGGGTGGTGTGGTGTGTTATCATGTTGTTGT 318
QY 81 ProProArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArgArgPro 100
DB 319 CCACCTAGAAATTCACCTACTGAGCGACACCTACAGTTGAATCGATCGAAGACGGCCC 378
QY 101 GlyGlnSerArgIleThrSerThrArgLysGluThrAspThrCysLysIleLeuSerGly 120
DB 379 CGACAGAGCAGAAATACCTCCACAGGAGGAGACTGATACATGCAAAATTCGTGAGGG 438
QY 121 ThrHisGluGlyValThrThrGlyThrProIleLeuValIleValProAsnThrAspGln 140
DB 439 ACACATGAAGGGGTGACTACTGGAACGCCAATTCCTTTATTTGTTCCCAACACACATCAA 498
QY 141 IleGlySerAspHisArgGluIleAlaAsnValTyrArgProSerHisAlaAspAlaThr 160
DB 499 ATAGGCAGTGTATCAGCGTGAATAGCCATGTGTACCGACCTCTCATGACAGCGCACT 558
QY 161 TyrAspPheLysTyrGlyValArgAlaValGlnGlyGlyArgSerSerGlyArgLys 180
DB 559 TATGACTTCAAGTACCGTGTAGAGCTGTACAGGGAGGTGGGAGGTCTCTCGGGCAGAAA 618

QY 181 ThrValGlyArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysLeuLysCysGly 200
DB 619 ACCGTTGGAAGGGTGGCTGCAGGGGCCCTCCCAAGAAAATTTTAAGCTCAAAATGGGA 678
QY 201 LeuGluIleLeuSerPheValSerLysValHisGlnValValLeuProGluAspAlaVal 220
DB 679 TTAGAGATCTTGTGTTGTTTCCAAAGTGCATCAGGTTGTGCTCCCAAGACGCGGTT 738
QY 221 AspTyrGlySerValThrLeuGluGlnIleGluSerAsnIleValArgCysProAspPro 240
DB 739 GATTATGGGCTGTAACTTTTGGAAACAGATAGAGACCAACATCGTTAGATGTCTCTGATCCA 798
QY 241 GluTyrAlaGluLysMetIleAspAlaIleAspArgValArgValArgGlyAspSerVal 260
DB 799 GAGTACCAGAGAAGATGATAGCGCAATGCACAGAGTACAGTTCCAGGGGATTCGGTC 858
QY 261 GlyGlyValIleThrCysValAlaArgAsnValProArgGlyLeuGlySerProValPhe 280
DB 859 GGTGAGTGTATCATCGTCTGCTAGAAACGTTCTCTCGGGCTCGGTTCTCTGTGTTC 918
QY 281 AspLysLeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPhe 300
DB 919 GACAAGCTCGAATCCGAACCTGGCAAAAGCTATGCTTTCTATTCTCGGACCAACGGGTT 978
QY 301 GluIleGlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPhe 320
DB 979 GAGATTGGCAGCGGATTCGCCGGGACCGACTTGCAGAAAGTAGACATATATGATGATTT 1038
QY 321 TyrMetAspLysAlaGlySerValArgThrArgThrAsnArgSerGlyValGlnGly 340
DB 1039 TATATGATTAAGGCTGGAAGTGTGCAGGACACGGGACTAATCGCTCGGGTGTGTGCAGGA 1098
QY 341 GlyIleSerAsnValGluIleValHisPheLysValAlaPheLysProThrProSerIle 360
DB 1099 GGGATATCGAATGTTGAGATTGTGCACTTCAAGTTGCTTTTAAAGCCGACCACTATCT 1158
QY 361 GlyValLysGlnAsnThrValSerArgGluArgGlnAsnValGluLeuLeuAlaArgGly 380
DB 1159 GGGGTGAACAGAACACCGTGTCAAGGAGCGCTCAGAACGTTGAGCTTCTAGCAGAGGG 1218
QY 381 ArgHisAspProCysValAlaProArgAlaValProValValGluSerMetAlaAlaLeu 400
DB 1219 CGCCATGACCCATCGCTCGCCCTCGAGTGTCTCTGTGGTGAATCCATGGCCGCGTTG 1278
QY 401 ValLeuMetAspGlnLeuMetAlaHisValAlaGlnCysGluMetPheAlaLeuAsnThr 420
DB 1279 GTCTCTATGGACCACTGATGGCGCACGTGGCTCAGTGCAGAGATGTTTCGCGCTCAACT 1338
QY 421 AlaLeuGlnGluProValGlySerPhe 429
DB 1339 GCATTTCAAGAACCACTGGCTCTTTTC 1365

RESULT 2

AAZ50435
ID AA250435 standard; cDNA; 1635 BP.
XX
AC AA250435;
XX
DT 18-MAY-2000 (first entry)
XX
DE Corn clone Chpc24.pk0002.h1:his encoding chorismate synthase.
XX
KW Corn cDNA clone chpc24.pk0002.h1:his; chorismate synthase;
KW immunological screening; herbicide resistance; antibody; gene mapping;
KW corn; ss.
XX
OS Zea mays.
XX
FH Location/Qualifiers
FT 52..1374
FT /*tag= a
FT /product= "Chorismate synthase"
FT

KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 200SEP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0128845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0132866P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137502P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147152P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147935P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.


```

PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 9,98e-125 Length: 1595
Score: 1461.00 Matches: 276
Percent Similarity: 81.91% Conservative: 59
Best Local Similarity: 67.48% Mismatches: 62
Query Match: 66.41% Indels: 12
DB: 3 Gaps: 2

US-10-677-179-8 (1-429) x AAC40227 (1-1595)
QY 30 ProAlaSerSerSerLeuArgPheAlaValHisArg----- 41
DB 176 CCGGTTCTTCTCTCTCCCTCGGAGCTCCGTCGTCCTTCTTCCCGCGCTTCAGATC 235
QY 42 ---CysArgThr-----AlaArgLeuGluValLysAlaSerGlyAsnThrPheGly 57
DB 236 TCTCTCCGTACCAACCGAGGAAGAACTTCCAGATACAAAGCTACTGGAAGTTTCATGGG 295
QY 58 AsnTyrPheGlnValAlaThrTyrGlyGluSerHisGlyGlyValGlyCysValIle 77
DB 296 ACTCATTTTCGAGTTTCAACTTTTCGAGAATCACATGGAGGAGGAGTTGTTGTATCAT 355
QY 78 SerGlyCysProProArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArg 97
DB 356 GATGGTTGTCTCTCTCGTATTCACCTTACTGAATCTGATTTGCAATTCGATTCGATAGA 415
QY 98 ArgArgProGlyGlnSerArgIleThrSerThrArgLysGluThrAspThrCysLysIle 117

```

```

DB 416 AGGAGACCTGGTCAGAGCAGGATCACAACCTCTAGAAAGGAGACTGATCTTCCGGATA 475
QY 118 LeuSerGlyThrHisGluGlyValThrThrGlyThrProIleLeuValIleValProAsn 137
DB 476 TCTTCTGGAGTCTCCGAAAGGAATGACGACAGGAACACCTATCCATGTGTGTGGCTAAC 535
QY 138 ThrAspGlnIleGlySerAspHisArgGluIleAlaAsnValTyrArgProSerHisAla 157
DB 536 ACAGATCAGAGAGGACTTGAATACAGTGAATGTCCGTTGCTTATAGACCATCGCATGCT 595
QY 158 AspAlaThrTyrAspPheLysTyrGlyValArgAlaValGlnGlyGlyValArgSerSer 177
DB 596 GATGCAACTTATGACATGAAGTATGTGTGTGCAGATCAGTGCAGGCTGGAGGAAGATCTTCA 655
QY 178 GlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysLeu 197
DB 656 GCTAGAGAGACCATTTGGAAGAGTTGCTCTCTGGAGCTTTGGCCAGAAAAATTTTGAAGCAA 715
QY 198 LysCysGlyLeuGluIleLeuSerPheValSerLysValHisGlnValValLeuProGlu 217
DB 716 TTTGAGGAACCTGAGATCTTGCCTATGTCTCGAAGTTCCACATGTTGTACTTCCAGAA 775
QY 218 AspAlaValAspTyrGlySerValThrLeuGluGlnIleGluSerAsnIleValArgCys 237
DB 776 GAATTTGGTAGACACGACAGAAATTTAACACTCGACACAGATAGAAAAATAACATGTGCAGATGC 835
QY 238 ProAspProGluTyrAlaGluLysMetIleAspAlaIleAspArgValArgValArgGly 257
DB 836 CTTAATCCGAGTATGCGAAAAAGATGATAGTGCATGATGATGCTGTCTCAGGACAAAAGGG 895
QY 258 AspSerValGlyGlyValIleThrCysValAlaArgAsnValProArgGlyLeuGlySer 277
DB 896 AACTCTGTGTGGTGGTGTGTGCCTGCATGTTCCGAATGTCTCCACGTGGGGCTTGGTACA 955
QY 278 ProValPheAspLysLeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSer 297
DB 956 CCGGTTTTCGATAACTTGAAGCAGAACTGGCAAAAGCTTGTATGTCGTACTTCCACACA 1015
QY 298 AsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsn 317
DB 1016 AAGGGATTTGAGTTTGGAAAGCGGCTTTGCAGGTACCTTTTGTACCTGGTCTTGAACACAAT 1075
QY 318 AspGluPheTyrMetAspLysAlaGlySerValArgThrArgThrAsnArgSerGlyGly 337
DB 1076 GATGAGTTCTATACCGATGAAATGGAAGAATACGTACCAAGCAAAACCGATCTGGTGA 1135
QY 338 ValGlnGlyGlyIleSerAsnValGluIleValHisPheLysValAlaPheLysProThr 357
DB 1136 ATTCAGGAGGGAATCTCAATGGTGAATATAAACAATGAGATGAGTAGCTTCAAGCCACA 1195
QY 358 ProSerIleGlyValLysGlnAsnThrValSerArgGluArgGlnAsnValGluLeuLeu 377
DB 1196 TCAACAATTTGGAAGGAAGCAGAAATACGGTAACCAAGAGAGGTAGAAACCGAAATGAT 1255
QY 378 AlaArgGlyArgHisAspProCysValAlaProArgAlaValProValValGluSerMet 397
DB 1256 GCGGCTGTCTCATGATCCTTGTGTGTTGTTCTCTGAGCTGTGCCAATGGTGAATCAATG 1315
QY 398 AlaAlaLeuValLeuMetAspGlnLeuMetAlaHisValAlaGlnCysGluMetPheAla 417
DB 1316 GTGGCTTTAGTTCTTGTGGATCAATTTGATGGCGCAATATGCGCAATGTTGTTTCCA 1375
QY 418 LeuAsnThrAlaLeuGlnGluProVal 426
DB 1376 ATAAATCCAGAGTTGCGAGAACCTCTC 1402
DB RESULT 4
ADN73862
ID ADN73862 standard; cDNA; 1311 BP.
XX
AC ADN73862;
XX

```


OS Bacteria.
 XX US2003233675-A1.
 XX 18-DEC-2003.
 XX 20-FEB-2003; 2003US-00369493.
 XX 21-FEB-2002; 2002US-0360039P.
 XX (CAOY/) CAO Y.
 XX (HINK/) HINKLE G J.
 XX (SLAT/) SLATER S C.
 XX (CHEN/) CHEN X.
 XX (GOLD/) GOLDMAN B S.
 XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX Claim 1; SEQ ID NO 26364; 122pp; English.
 XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polynucleotide used in
 CC the scope of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX SQ Sequence 1089 BP; 278 A; 245 C; 314 G; 252 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: Length: 1089
 Score: 1191.50 Matches: 225
 Percent Similarity: 77.90% Conservative: 57
 Best Local Similarity: 62.15% Mismatches: 77
 Query Match: 54.16% Indels: 3
 DB: 13 Gaps: 1
 US-10-677-179-8 (1-429) x ADSA7934 (1-1089)
 QY 53 GlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGlyGly 72
 DB 4 GGAACACGCTTTGGCAGTTTATTCGATTACCACTCCGTCGAGTCCCTCGGCGCGC 63
 QY 73 ValGlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAspLeuGln 92
 DB 64 GTGGGGTAATATTGATGGTTCCTCCCGATTGGAATTTCCCGAGGAAATCCAA 123
 QY 93 ValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLysGluThr 112

DB 124 GTAGACGTAGACCGCGCGGCCAGCAGACAGCAAAATTTACCAACCCCTCGTAAGGAAGCA 183
 QY 113 AspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThrProIleLeu 132
 DB 184 GATCAATCGAAATTTCTCTGCGGTGTTGAGGGGTAAACACTCTGGGTACACCGATCGCC 243
 QY 133 ValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnValTyr 152
 DB 244 ATTTGGTGGGAATAAGATGCCCGTCCAGGATTACAATGAGATGCGGTGAATAT 303
 QY 153 ArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValAlaGlnGly 172
 DB 304 CGTCTTCCACGCGCATGCCACCTATGAAGCGAAATATGGCATTCGCAATTTGGCAAGGA 363
 QY 173 GlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuProLys 192
 DB 364 GGAGTGGCTCTTCTGCTAGGGAACCATTTGGCCGGGTGGCCGAGGGCGCATCGCAAA 423
 QY 193 LysIleLeuLysLysCysGlyLeuLysGlyLeuLysSerPheValSerLysValHisGln 212
 DB 424 AAGATTCTGGCCCGAGTTTAAACGGGGTGAATTTGGCTATGTGAAGATATCCAAGAC 483
 QY 213 ValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGluSer 232
 DB 484 ATT-----GAGGCGACGGTGCACAGCAATACGGTGACCTTGGAGCAAGTGAAGAGC 534
 QY 233 AsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIleAspArg 252
 DB 535 AATATTGTGCTGTGTCGCGATGAAGATGTGCCGAAATAATGATTGAACGCATTCACCAAG 594
 QY 253 ValArgValArgGlyAspSerValGlyGlyValIleThrCysValAlaArgAsnValPro 272
 DB 595 GTGTACACACAAAGACTCCATTTGGCGGGGTGGTGGATGTGCCATCGTAAACGCTCC 654
 QY 273 ArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMetLeu 292
 DB 655 AAGGGTGTGGGGAAACAGTGTTCATTAAGTTGGAAGCGGATTTGGCCAAAGCAATGATG 714
 QY 293 SerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeuThr 312
 DB 715 TCCCTCCCGCTTACCAAGGGGTGTGAGTTTGCTCTGCTGTTTGGCGGTACTTTACTGACT 774
 QY 313 GlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThrArgThr 332
 DB 775 GGTAGTCAACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 834
 QY 333 AsnArgSerGlyValGlnGlyGlyIleSerAsnValGluIleValHisPheLysVal 352
 DB 835 AACCGTTCTGGGGGGGTTTCAGGGAGGCATCAGCAACGGCGAACCCATTTATTATGCGCATA 894
 QY 353 AlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGluArgGln 372
 DB 895 GCATTTAAACCTACTGCCACCATTTGGTCAAGACGCAAAACCGTTAGTAACATTTGGGAG 954
 QY 373 AsnValGluLeuAlaArgGlyArgHisAspProCysValAlaProArgAlaValPro 392
 DB 955 GAAACACCTTAGCGCTAAGGTTCGCGCATGATCCCTCGGTTCCTCCCGCGGTGTGCC 1014
 QY 393 ValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisValAlaGln 412
 DB 1015 ATGGTGAAGCGATGGCTGCATTAGTACTTTGTGTGACCATTTTGTTCGATTTTCCAGCC 1074
 QY 413 CysGlu 414
 DB 1075 TGTAAA 1080
 RESULT 8
 ADT44117
 ID ADT44117 standard; cDNA; 1086 BP.
 XX AC ADT44117;
 XX DT 02-DEC-2004 (first entry)

XX Bacterial polynucleotide #18868.
DE Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.
XX Bacteria.
OS
XX
XX US2003233675-A1.
PN
XX
XX 18-DEC-2003.
PD
XX
XX 20-FEB-2003; 2003US-00369493.
PF
XX
XX 21-FEB-2002; 2002US-0360039P.
PR
XX
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI WPI; 2004-061375/06.
XX
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 42555; 122pp; English.
PS
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX
SQ Sequence 1086 BP; 308 A; 216 C; 290 G; 272 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,2e-99 Length: 1086
Score: 1186.50 Matches: 220
Percent Similarity: 78.79% Conservative: 66
Best Local Similarity: 60.61% Mismatches: 74
Query Match: 53.93% Indels: 3
DB: 13 Gaps: 1

US-10-677-179-8 (1-429) x ADT44117 (1-1086)

QY 53 GlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGlyGly 72
DB 4 GGC AATACTTTTGGACATCTGTTTCGGATCACTACTTTTGGCGAGTCACACGGTGGAGGC 63
QY 73 ValGlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAspLeuGln 92
DB 64 GTGGGGTGTGATTGACGGCTGCTCCGCTACTCGAAATTTCTCTCGAGGAATTCAA 123
QY 93 ValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLysGluThr 112
DB 124 CTGGAGTTAGACCGACGACCGACAGAGTAAGATTACACGCCCTCGTAAGGAAGCA 183
QY 113 AspThrCysLysIleLeuSerGlyThrHisGluGlyValThrGlyThrProIleLeu 132
DB 184 GACACCTGTGAGATTTTGTCTGGGGTATATGAAGGCAAACTCTGGGAACCCCAATATCA 243
QY 133 ValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnValTyr 152
DB 244 ATTTTGTGCGAATAAAGATACCTCGTCCCAAGATTATGATGAGATGGCGCAAAAGTAT 303
QY 153 ArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGlnGly 172
DB 304 CGTCTTCCCATGCTGATGCAACTTATGATGCTAAATATGGGATTTCGTAATTGGCAAGGT 363
QY 173 GlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuProLys 192
DB 364 GCGGTAGATCCTCAGCGCGGAGCAATTTGTTAGATAGCCGAGGTGCGATCGCTAAA 423
QY 193 LysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysValHisGln 212
DB 424 AAAATTCCTCGGCAAGTTGCTAATAGTAGAAGTTATAGCTACGTCAAAAGTATCAAAAGAC 483
QY 213 ValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGluSer 232
DB 484 TTG-----GAAGGTGTTGTAGATCCCAATACCGTCACCTAGACCAATGCGAAAGC 534
QY 233 AsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIleAspArg 252
DB 535 AATATCGTTCCTGCTCTGATGGCGAATGCGACACCGCATGATGTAATTAATCGAACAA 594
QY 253 ValArgValArgGlyAspSerValGlyValIleThrCysValAlaAlaArgAsnValPro 272
DB 595 ACAGGTAGACAAGGTGATTCTATCGGGGTGTAGTGAATCGGTAGCGCGCAATGTTCTCT 654
QY 273 ArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMetLeu 292
DB 655 AAAGCGTTGGGTGAACCTGTATTTCGATAAGTTGGAAGCAGATATCGCCAAAGCTGTGTATG 714
QY 293 SerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeuThr 312
DB 715 TCTTTACCTGCTAGCAAAAGTTTGAATCGTTTCAGGTTTTTCCGGACTCTCGTAACA 774
QY 313 GlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThrArgThr 332
DB 775 GGAATTGAGCATACGACGAGATATTATATTGATAAAGCGGTGAGATCGCACAGTAAC 834
QY 333 AsnArgSerGlyGlyValGlnGlyIleSerAsnValGluIleValHisPheLysVal 352
DB 835 AACCGTTCCGGTGGGATTCGAAGCGGTATTGCCAACCGAGAGAAATATCATTTTGGCGATT 894
QY 353 AlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGluArgGln 372
DB 895 GCATTTAAACCCACGCCCGCATTTAGAAAAGAGCAAAAACAGTTACTCTGTAAGGCGCA 954
QY 373 AsnValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaProArgAlaValPro 392
DB 955 GAAACCTTATTAGCGCCCAAGGTAGACATGATCCCTGTTCTTCTACCCCGTCAGTCCCG 1014
QY 393 ValValGluSerMetAlaLeuValLeuMetAspGlnLeuMetAlaHisValAlaGln 412
DB 1015 ATGTTAGGAAGCAATGTCGCGTGTGTTACTGTCGCGATCATTTTGTACGCCATCATGACAG 1074

```
QY 413 CysGluMet 415
Db 1075 TGTAAGGTG 1083

RESULT 9
ADT45459
ID ADT45459 standard; cDNA; 1086 BP.
XX
XX AC
XX ADT45459;
XX
XX DT 02-DEC-2004. (first entry)
XX
XX DE Bacterial polynucleotide #20210.
XX
XX KW Recombinant DNA construct; transformed plant; improved plant property;
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX pathogen tolerance; pest tolerance; plant disease resistance;
XX cell cycle pathway modification; plant growth regulator;
XX homologous recombination; seed oil yield; protein yield; carbohydrate;
XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX bacterial polynucleotide; gene; ss.
XX
XX OS Bacteria.
XX
XX PN US2003233675-A1.
XX
XX PD 18-DEC-2003.
XX
XX PF 20-FEB-2003; 2003US-00369493.
XX
XX PR 21-FEB-2002; 2002US-0360039P.
XX
XX (CAO Y) CAO Y.
XX (HINKLE G J.
XX (SLATER S C.
XX (CHEN X.
XX (GOLDMAN B S.
XX
XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
XX PT New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.
XX
XX PS Claim 1; SEQ ID NO 43897; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
XX promoter functional in a plant cell, where the promoter is positioned to
XX provide for expression of a polynucleotide encoding a polypeptide from a
XX microbial source. The invention also relates to a transformed plant
XX comprising the recombinant DNA construct and a method of producing a
XX transformed plant having an improved property. The plant is a crop plant
XX such as maize or soybean. The method of producing a transformed plant
XX having an improved property comprises transforming a plant with the
XX recombinant DNA construct and growing the transformed plant, where the
XX polynucleotide or polypeptide is useful for improving plant properties.
XX The recombinant DNA construct is useful for producing plants with
XX improved plant properties, e.g. improved cold, heat or drought tolerance,
XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX increased resistance to plant disease, better growth rate by modification
XX of the cell cycle pathway with plant growth regulators, increased rate of
XX homologous recombination, modified seed oil or protein yield and/or
XX content, improved yield by modification of carbohydrate, nitrogen or by
XX phosphorus use and/or uptake, by modification of photosynthesis or by
XX providing improved plant growth and development under at least one stress
XX condition, improved lignin production or improved galactomannan
XX production. This sequence represents a bacterial polynucleotide used in
XX the scope of the invention. Note: The sequence data for this patent did
XX not form part of the printed specification but was obtained in electronic
XX format from USPTO at seqdata.uspto.gov/sequence.html.
XX
```

Sequence 1086 BP; 319 A; 201 C; 290 G; 276 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.79e-99	Length:	1086
Score:	1182.50	Matches:	220
Percent Similarity:	78.51%	Conservative:	65
Best Local Similarity:	60.61%	Mismatches:	75
Query Match:	53.75%	Indels:	3
DB:	13	Gaps:	1

US-10-677-179-8 (1-429) x ADT45459 (1-1086)

QY 53 GlyAsnThrPheGlyAsnTyPheGlnValAlaThrTyGlyGluSerHisGlyGly 72

Db 4 GGCATATTTTGGTCAATTTATTCGCATAGTACTTTTGGCGAGTCTCACGGCGGT 63

QY 73 ValGlyCysValIleSerGlyCysProProArgIleProLeuThrGluAlaAspLeuGln 92

Db 64 GTGGGGTGTGATTTGATGATGTTCTCCACAACTAGAAAATTCGGCAGAGAAATTCAG 123

QY 93 ValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgGlySerGluThr 112

Db 124 GTAGAACTAGATAGAAAGCGCGGCAAAATTAACGACTCCCGCAAGAGCT 183

QY 113 AspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThrProIleLeu 132

Db 184 GATACCTCGGAGATTTATCTCAGGAGTATTTGAGGCAAAACGCTGGGACCCCTATACG 243

QY 133 ValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnValTyr 152

Db 244 ATTTGGTACGTAATCAAGATACTCGTCCCAAGATTTATGACGAGATGCGACAGAGTAT 303

QY 153 ArgProSerHisAlaAspAlaThrTyRaspPheLysTyGlyValArgAlaValGlnGly 172

Db 304 CGGCTTCTCAGCGGATGCAACCTATGATGCAAAATATGCGCATTCGCAATTTGGCAAGT 363

QY 173 GlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyValAlaLeuProLys 192

Db 364 GGGGTAGTGTCTCAGCAGCTGAGACAATCGGAGAGTAGCAGCAGGTGCGATGCTAAA 423

QY 193 LysIleLeuLysLeuLysCysGlyGluLeuLeuSerPheValSerLysValHisGln 212

Db 424 AAAATTCCTCGTCAAGTTGCCAATGTCGAAATATTTGCTTACGTTAAGCGGATTAAGAC 483

QY 213 ValValLeuProGluAspAlaValAspTyRaspGlySerValThrGluGlnIleGluSer 232

Db 484 TTG-----GAAGGTGTAGTCGATCCTAATCTACTGTAACCTTAGAACAAGTAAAGC 534

QY 233 AsnIleValArgCysProAspProGlyTyAlaGluLysMetIleAspAlaIleAspArg 252

Db 535 AATATCGTTCGCTGTCGCGATGCTGATGTCGCGATCGCATGCGATGATTAATGATGAGCA 594

QY 253 ValArgValArgGlyAspSerValGlyValIleThrCysValAlaAlaArgAsnValPro 272

Db 595 ATAGTGACACAAGGGGATTTCTATCGCGGTGTCGTAGAATGTGTGGCGCAATGTACCG 654

QY 273 ArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMetLeu 292

Db 655 AAAGTTTGGGCGAACCACTGATTTGATAAATTTAGAAAGCTGATATCGCTAAGGGGTGATG 714

QY 293 SerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeuThr 312

Db 715 TCTCGCTCTAGCAAGGCTTTGAAATTTGTTCCGGTTTTCGGGACCGCTACTACG 774

QY 313 GlySerGluHisAsnAspGluPheTyMetAspLysAlaGlySerValArgThrArgThr 332

Db 775 GGAATTTAGACATAACGACGAATTTTATATCGATCAAAATGATGAAATTCGTACAGTAAACA 834

QY 333 AsnArgSerGlyGlyValGlnGlyGlyLeuSerAsnValGluIleValHisPheLysVal 352

Db 835 AACCGTTCTGGCGGTATTCAAGCGCGGATTTCTAACCGCGGAAATATCATTTTGGCGATT 894

QY 353 AlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGluArgGln 372

pathogen tolerance; pest tolerance; plant disease resistance; cell cycle pathway modification; plant growth regulator; homologous recombination; seed oil yield; protein yield; carbohydrate; nitrogen; phosphorus; photosynthesis; lignin; galactomannan; bacterial polynucleotide; gene; ss.

Bacteria.

US2003233675-A1.

18-DEC-2003

20-FEB-2003: 2003US-00369493

21-FEB-2002: 2002US-0360039P

(CAOY/) CAO Y.

(HINK//) HINKLE G J.

(SLAT/) SLATER S C.

(CHEN/) CHEN X.

(GOLD/) GOLDMAN B S.

Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

WPI; 2004-061375/06.

New recombinant DNA construct comprising a promoter positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source, useful for producing plants with improved properties.

Claim 1; SEQ ID NO 44740; 122pp; English.

The invention relates to a recombinant DNA construct comprising a promoter functional in a plant cell, where the promoter is positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source. The invention also relates to a transformed plant comprising the recombinant DNA construct and a method of producing a transformed plant having an improved property. The plant is a crop plant such as maize or soybean. The method of producing a transformed plant having an improved property comprises transforming a plant with the recombinant DNA construct and growing the transformed plant, where the polynucleotide or polypeptide is useful for improving plant properties. The recombinant DNA construct is useful for producing plants with improved plant properties, e.g. improved cold, heat or drought tolerance, tolerance to herbicides, extreme osmotic conditions, pathogens or pests, increased resistance to plant disease, better growth rate by modification of the cell cycle pathway with plant growth regulators, increased rate of homologous recombination, modified seed oil or protein yield and/or content, improved yield by modification of carbohydrate, nitrogen or phosphorus use and/or uptake, by modification of photosynthesis or by providing improved plant growth and development under at least one stress condition, improved lignin production or improved galactomanan production. This sequence represents a bacterial polynucleotide used in the scope of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html.

Sequence 1080 BP; 204 A; 331 C; 369 G; 176 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.09e-85	Length:	1080
Score:	1035.00	Matches:	200
Percent Similarity:	72.65%	Conservative:	63
Best Local Similarity:	55.25%	Mismatches:	95
Query Match:	47.05%	Indels:	4
DB:	13	Gaps:	2

US-10-677-179-8 (1-429) x ADT46302 (1-1080)

Ov 53 GlvAsnThrPheGlvAsnTyrPheGlnValAlaThrTyrGlvGlvSerH⁺SGlvGlvGlv 72

Db
4 GGCAGCAGCTTCGGCGGATCTGTTCGGATCAGCACCTTCGGTGAGTCCCA CGGCGGTGGC 63

73 ValGlyCysValIleSerGlyCysProArgGileProLeuThrGluAlaAspLeuGln 92
Db : : : : :
64 GTGGGGTGATCGTGGAGGGCTGCCCTCCACGCCTGGATCTGGATGTTCGACGCATCCAG 123

93 ValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLysGluThr 112
Db : : : : :
124 GCGGAGTTGGATCGGGACGCGCGGGTTCGATGGCGCAGACCACCCACCACCAAGAAGCC 183

113 AspThrCysLysIleLeuSerGlyThrHisgluglyValThr---GlyThrProIle 131
Db : : : : :
184 GATCAGGTGGAGGTCTCAGCGCCCTTCGATGGCGCAGACCATCCCTGGGAAACACCGATC 243

132 LeuValIleValProLeuThrAspGlnIleGlySerAspHisArgGluIleAlaAsnVal 151
Db : : : : :
244 GCCATTGTGTGGCACCAAGGACACGCGCCCCGGGACTACAAGGACATGGCGGTGGCG 303

152 TyrArgProSerHisalaAspAlaThrTyrAspPheIysTyrGlyValArgAlaValGln 171
Db : : : : :
304 TTTTGGCCCTTCTCATGCCGATGCCACTACGAGTGAAATACCGCGTGCAGGCGCCGAGT 363

172 GlyGlyGlyArgSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuPro 191
Db : : : : :
364 GTTGGCGGGCGGCCTCAGCCCGGAACAAATTTGGCGGTGGCGCGGTGCATATGCC 423

192 LysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValserIysValHis 211
Db : : : : :
424 AAGCAACTGTGCAAAAGCGCGCAGGCACGAGAGGTCTTGGCCTGGGTCAAACGCATCCAC 483

212 GlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluIleGlu 231
Db : : : : :
484 ACCATC-----GAGGCGAATGTGGATCCCAGCGCGGTAACTGGTGGACGCATCGAG 534

232 SerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIleAsp 251
Db : : : : :
535 AGCAACATGTTTCGTCGCCGATGCTCCACAGCGCTCAGATGGTGGAACGGATCGAG 594

252 ArgValArgValArgGlyAspSerValGlyValIleThrCysValAlaArgAsnVal 271
Db : : : : :
595 CGCATCGGGCGCAGCGGACTCATGTGTGGGTGATCGAATGCGTGGTGGCTTAACCT 654

272 ProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluIleAlaLysAlaMet 291
Db : : : : :
655 GCTCCGGGGTGGGTATGCTGTGTTCGACAAAGCTGGAGCGCATCTGGCCAAAGGCGGTG 714

292 LeuSerIleProAlaserAsnGlyPheGluIleGlySerClyPhealaglyThrAspLeu 311
Db : : : : :
715 ATGTCTCTCCCCCCCAACGAGGCTTTGAGATCGGTCTCTGGCTTCAGCGGAAACCTCGCTC 774

312 ThrGlySerGluHisAsnAspGluPheTyMetAspLysAlaGlySerValArgThrArg 331
Db : : : : :
775 AAGGGCAGCGAGCACACGATGCTTTTGTGCCACCGATGATGCGGTGCAGACGGCC 834

332 ThrAsnArgSerGlyGlyValGlnGlyIleSeraSnValGluIleValHisPheLys 351
Db : : : : :
835 ACCAAACAATCAGTGGCATCCAGGCGGCATCAGCAACGCGCAGCGCATTTGTATCGG 894

352 ValAlaPheIysProThrProSerIleGlyValLysGlnAsnThrValSerArgGluArg 371
Db : : : : :
895 GTGGCGTTCAAGCCCAACCGCCAGTCCGCAAGAACACGACGACCAATGATTCGACGGC 954

372 GlnAsnValGluLeuAlaArgGlyArgHisAspProCysValAlaProArgAlaVal 391
Db : : : : :
955 AAGGCCACACCTTGGCGGCAGGGTTCGACACGACCCCTTGTGTGTCTGCGAGGGCGGTG 1014

392 ProValValGluSerMetAlaAlaLeuValLeuMetAspDlnLeuMetAlaHisValAla 411
Db : : : : :
1015 CCGATGTGGAGCGCATGTTGGCGCTGGTGTCTGCGGATCACCTGTCTGCGGACGAGGT 1074

412 GlnCys 413
Db : : : : :
1075 CAGTGC 1080

RESULT 12

[illegible]

Db 955 GTTACTGTAGCGGTAAAGCGCCGATGACCCATGCGTGTGTTCCAGAGCGGTACCGGATT 1014

QY 394 ValGluSerMetAlaAlaLeuValLeuMetAsp 404

Db 1015 GTTGAAGCAATGCGGCAATTGTTATGTGTGGAT 1047

RESULT 13

ADT47171

ID ADT47171 standard; cDNA; 1131 BP.

XX AC ADT47171;

XX DT 02-DEC-2004 (first entry)

XX DE Bacterial polynucleotide #21922.

XX KW Recombinant DNA construct; transformed plant; improved plant property; cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis; pathogen tolerance; pest tolerance; plant disease resistance; cell cycle pathway modification; plant growth regulator; homologous recombination; seed oil yield; protein yield; carbohydrate; nitrogen; phosphorus; photosynthesis; lignin; galactomannan; bacterial polynucleotide; gene; ss.

XX OS Bacteria.

XX PN US2003233675-A1.

XX PD 18-DEC-2003.

XX PF 20-FEB-2003; 2003US-00369493.

XX PR 21-FEB-2002; 2002US-0360039P.

XX PA (CAOY/) CAO Y.

XX PA (HINK/) HINKLE G J.

XX PA (SLAT/) SLATER S C.

XX PA (CHEN/) CHEN X.

XX PA (GOLD/) GOLDMAN B S.

XX QY Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX PT New recombinant DNA construct comprising a promoter positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source, useful for producing plants with improved properties.

XX PS Claim 1; SEQ ID NO 45609; 122pp; English.

XX CC The invention relates to a recombinant DNA construct comprising a promoter functional in a plant cell, where the promoter is positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source. The invention also relates to a transformed plant comprising the recombinant DNA construct and a method of producing a transformed plant having an improved property. The plant is a crop plant such as maize or soybean. The method of producing a transformed plant having an improved property comprises transforming a plant with the recombinant DNA construct and growing the transformed plant, where the polynucleotide or polypeptide is useful for improving plant properties. The recombinant DNA construct is useful for producing plants with improved plant properties, e.g. improved cold, heat or drought tolerance, tolerance to herbicides, extreme osmotic conditions, pathogens or pests, increased resistance to plant disease, better growth rate by modification of the cell cycle pathway with plant growth regulators, increased rate of homologous recombination, modified seed oil or protein yield and/or content, improved yield by modification of carbohydrate, nitrogen or phosphorus use and/or uptake, by modification of photosynthesis or by providing improved plant growth and development under at least one stress condition, improved lignin production or improved galactomannan production. This sequence represents a bacterial polynucleotide used in the scope of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic

CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 1131 BP; 312 A; 267 C; 272 G; 280 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.01e-76 Length: 1131

Score: 938.00 Matches: 183

Percent Similarity: 68.12% Conservative: 67

Best Local Similarity: 49.86% Mismatches: 109

Query Match: 42.64% Indels: 8

DB: 13 Gaps: 3

US-10-677-179-8 (1-429) x ADT47171 (1-1131)

QY 54 AsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGlyGlyVal 73

Db 4 TCACGCTTTGGGAAACTGTTCCGCTCACCATATGGTGAATGCATTTGTAAGTGTGTC 63

QY 74 GlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAspLeuVal 93

Db 64 GGTTCATGTGTCAGCGGTCTTCAGGAATGTCTTAACCAAGTGCATTCAGCCA 123

QY 94 GluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLysGluThrAsp 113

Db 124 CAATTGACCAGAAAGACCGGTCATCTAAGCTATCCACCCCTAGACACGAAAGGAT 183

QY 114 ThrCysLysIleLeuSerGlyThrHisGluGlyValThrGlyThrProIleLeuVal 133

Db 184 AGAGTGGAAATCCAGTCCGCTACCGAGTTCCGCAAGACTCTAGGTACACCCATCGCATG 243

QY 134 IleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnValTyrArg 153

Db 244 ATGATCAAAACGAGACCAAGACCTCAGCACTACTCGCATCGACAGTTCCTTAGA 303

QY 154 ProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGlnGly 173

Db 304 CTTTCCCATGCGGACTTCAGCTACTCGGAAAGTAGCGGTATCAAGGCTCTCTCTGTTG 363

QY 174 GlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuProLys 193

Db 364 GGCAGAGCTTCTGCTAGAGAAACGATTCGCTTCAGGTTCAGGTGCATTCCTGAGAAG 423

QY 194 IleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysValHisGln 213

Db 424 TTTCTAGCTCAGAACTCTAATGTCAGATCGTAGCTTTGTGACACAATCGGGAAATC 483

QY 214 ValLeuProGluAspAlaValAsp-----TyrGlySerValThrLeu 227

Db 484 AAGATGAACAGAGACTCTTTGATCTCTGATCTTTCAGCATCTGTTGAACACCATCAGG 543

QY 228 GluGlnIleGluSer---AsnIleValArgCysProAspProGluTyrAlaGluLysMet 246

Db 544 GAAAGAGTGGACTCAATGGGTCTTATCAGATGTCAGACGCTCCGCTGCTGTTGATG 603

QY 247 IleAspAlaIleAspArgValArgGlyAspSerValGlyGlyValIleThrCys 266

Db 604 GTCAAGGAAATCGAAAGTACAGAGCAACAGGACTCTATCGGTGGTGTCTGCTACTGT 663

QY 267 ValAlaArgAsnValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGlu 286

Db 664 GTCTGAGAAACTTGCTTACCGTCTCGTTCGATCGCTTGTGACAAGTGGAGCCATG 723

QY 287 LeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPhe 306

Db 724 TTGGCTCATGTATGTGCTTCCATTCAGCATCCAGGGGTTTCGAAATTTGGCTCAGGTTT 783

QY 307 AlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPheTyrMetAspLys---Ala 325

Db 784 CAGGTGTCTCTGTCCAGGGTCCAGACCAATGACCCATTTTACTTTGAAGAAGAAACA 843

QY 326 GlySerValArgThrArgThrAsnArgSerGlyValGlnGlyValGlyIleSerAsnVal 345

Db 844 AACAGATTAAAGACAAAGACCAACCAATTCAGGTGTGTACAGGTGGTGTATCTCTTAATG 903

QY 346 GluIleValHisPheLysValAlaPheLysProThrProSerIleGlyValLysGlnAsn 365
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 904 GAGAACATCTATTTCTCTGTCCTCCATCAAGTCAGTGGCCACTATCTTCACAGAACAAAA 963
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 366 ThrValSerArgGluArgGlnAsnValGluLeuLeuAlaArgGlyValHisAspProCys 385
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 964 ACCGCCACTTACGATGGTGAAGAAGTATCTTAGCGGCTAAGGTAGACATGACCTGCT 1023
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 386 ValAlaProArgAlaValProValValGluSerMetAlaAlaLeuValLeuMetAspGln 405
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1024 GTCACTCCAAGAGCTATTCCTATTCTGGAAGCCACGCGCTCTGGTGTGGCTGACGCG 1083
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 406 LeuMetAlaHisValAlaGln 412
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1084 CTTTGTATCCAAAGCAGA 1104
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
RESULT 14
ADP98700
ID ADP98700 standard; DNA; 1137 BP.
XX
AC ADP98700;
XX
DT 23-SEP-2004 (first entry)
XX
DE C. albicans specific gene, orf6.1285, DNA sequence.
XX
KW Diploid fungal cell; allele; gene disruption cassette;
KW promoter replacement fragment; antifungal; fungicide; gene therapy;
KW infection; Candida albicans; gene; ds.
XX
OS Candida albicans.
XX
PN WO2004056965-A2.
XX
PD 08-JUL-2004.
XX
PF 19-DEC-2003; 2003WO-US040618.
XX
PR 19-DEC-2002; 2002US-0434832P.
XX
PA (ELIT-) ELITRA PHARM INC.
PA (ELIT-) ELITRA CANADA LTD.
XX
PI Roemer T, Jiang B, Boone C, Bussey H;
XX
XX WPI; 2004-500296/47.
DR P-PSDB; ADP98700.
XX
PT Constructing a strain of diploid fungal cells in which both alleles of a
PT gene are modified comprises modifying the alleles of a gene in the fungal
PT cells by recombination using a gene disruption cassette and a promoter
PT replacement fragment.
XX
PS Claim 33; SEQ ID NO 6185; 163pp; English.
XX
CC The invention relates to a novel method for constructing a strain of
CC diploid fungal cells in which both alleles of a gene are modified. The
CC method comprises modifying the alleles of a gene in diploid fungal cells
CC by recombination using a gene disruption cassette and a promoter
CC replacement fragment. The invention further comprises: assembling a
CC collection of diploid fungal cells each of which comprises modified
CC alleles of a different gene; a strain of diploid fungal cells comprising
CC modified alleles of a gene, where the first allele of the gene is
CC inactivated by a gene disruption cassette comprising a nucleotide
CC sequence encoding an expressible selectable marker; and the expression of
CC the second allele of the gene is regulated by a heterologous promoter
CC that is operably linked to the coding region of the second allele of the
CC gene, and where the gene encodes the polypeptide mentioned above; a
CC collection of diploid fungal strains comprising the diploid strains cited
CC above, where substantially all the different genes that encode the above
CC amino acid sequences are modified and are present in different diploid
CC strains in the collection; a nucleic acid molecule microarray comprising

CC nucleic acid molecules, where each nucleic acid molecule comprises a
CC nucleotide sequence that is hybridizable to a target nucleotide sequence
CC comprising any of the 310 nucleotide sequences listed in the
CC specification (ADP98516-ADP98825); identifying a gene that is essential
CC and/or pathogenicity of a fungus, that contributes to the virulence
CC of a diploid fungus to an antifungal agent; identifying an antifungal
CC agent that inhibits the growth of a diploid fungus, or a therapeutic
CC agent for treatment of a mammalian disease; correlating changes in the
CC levels of proteins or gene transcripts with the inhibition of growth or
CC proliferation of a diploid fungal cell; a purified or isolated nucleic
CC acid molecule comprising a nucleotide sequence encoding a gene product
CC required for proliferation of Candida albicans, where the gene product
CC consists of any of the above-mentioned amino acid sequences; a vector
CC comprising a promoter operably linked to the nucleic acid molecule cited
CC above; a host cell containing the vector; a purified or isolated
CC polypeptide comprising any of the 61 amino acid sequences given in the
CC specification (ADP96718-ADP96778); a fusion protein comprising a fragment
CC of a first polypeptide fused to a second polypeptide, the fragment
CC consisting of at least 6 consecutive residues of any of ADP98826-ADP99135
CC; producing a polypeptide; identifying a compound which modulates the
CC activity of a gene product encoded by a nucleic acid comprising any of
CC ADP98516-ADP98825; eliciting an immune response in an animal; a strain of
CC Candida albicans, where a first allele of a gene comprising any of
CC ADP98516-ADP98825 is inactive and a second allele of the gene is under
CC the control of a heterologous promoter; identifying a compound or binding
CC partner that binds to the polypeptide comprising any of ADP98826-
CC ADP99135, or its fragment; identifying a compound having the ability to
CC inhibit growth or proliferation of Candida albicans; inhibiting growth or
CC proliferation of Candida albicans cells; manufacturing an antimycotic
CC compound; treating an infection of a subject by Candida albicans;
CC preventing or containing contamination of an object by Candida albicans,
CC or for preventing or inhibiting formation on a surface of a biofilm
CC comprising Candida albicans; a pharmaceutical composition comprising a
CC therapeutic amount of an agent which reduces the activity or level of a
CC gene product encoded by a nucleic acid comprising any of ADP98516-
CC ADP98825 in a pharmaceutical carrier; an antibody preparation which binds
CC the polypeptide; methods for evaluating a compound against a target gene
CC product encoded by any of ADP98516-ADP98825; identifying an antimycotic
CC compound; a computer or a computer readable medium that comprises at
CC least one of the nucleotide sequences mentioned in the specification or
CC method assisted by a computer for identifying a putatively essential gene
CC of a fungus; and a protein array comprising proteins, where at least one
CC protein comprises an amino acid sequence or a portion of an amino acid
CC sequence selected from ADP98516-ADP98825. The novel methods and
CC compositions have fungicide activity. The compositions may be used in
CC gene therapy. The composition and methods are useful for drug screening
CC purposes or for diagnosing, preventing or treating infections associated
CC with Candida albicans. These may also be used for constructing strains
CC useful for identification and validation of gene products as effective
CC targets for therapeutic intervention, for identifying and validating gene
CC products as effective targets for therapeutic intervention, and for
CC collecting identified essential genes. This polynucleotide sequence
CC represents a Candida albicans fungal specific gene of the invention.
CC NOTE: This sequence was downloaded from an electronic sequence listing
CC provided on the WIPO website.

SQ Sequence 1137 BP; 374 A; 185 C; 251 G; 327 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.37e-76	Length:	1137
Score:	934.00	Matches:	188
Percent Similarity:	68.39%	Conservative:	63
Best Local Similarity:	51.23%	Mismatches:	108
Query Match:	42.45%	Indels:	8
DB:	12	Gaps:	4

US-10-677-179-8 (1-429) x ADP98700 (1-1137)

QY 54 AsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGlyGlyVal 73
: : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 4 TCGTCTTTGGACATTTATTTAAAGTTACTACATATGGGAATCACACTGTAATACAGTT 63

```
QY 74 GlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAspLeuGlnVal 93
DB 64 GGGTGTATTGTTGATGGTGTCCACCGAGGATGCTTAATCTGACGAGATATTCACCA 123
QY 94 GluLeuAspArgArgArgProGlyGlnSerArgIleThrSerThrArgLysGluThrAsp 113
DB 124 CAATTGACTCGTAGAGCTCCAGGTCAAAGTAAATTTATCTACCCCAAGAGATGAAAAGAT 183
QY 114 ThrCysIleLeuSerGlyThrHisGluGlyValThrGlyThrProIleLeuVal 133
DB 184 TTAGTTGAATAACAAGTGAACCTGAGATGGAATTAACCTTGGGTTCACCTATTGGAATG 243
QY 134 IleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnValTyrArg 153
DB 244 ATTGTGAGAAATAAAGATCATCTCTCTGCTGATTCAGTGAGACCGATTGTATCCAGA 303
QY 154 ProSerHisAlaAspAlaThrTyrAspPheLysThrGlyValArgAlaValGlnGly 173
DB 304 CCAAGTCATGCCGATTCGACCTATATACAGAAATACGGAACCAAAATCTAGTTCTGGTGT 363
QY 174 GlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGluLeuProLysLys 193
DB 364 GGTAGATCATCTCTAGAGAAATATTTGGTAGAGTTGCCGACGAGCCATTGCGAAGAAAG 423
QY 194 IleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysValHisGlnVal 213
DB 424 ATTTTGGCCAAAGTGAATAATGTTGAATTTGCTTTGTAGTCTATTGTGTGAATA 483
QY 214 ValLeu-----ProGluAspAlaValAsp-----TyrGlySerValThrLeu 227
DB 484 TCTATGAGTAAATCTCTCAAGATGCAAAATTCGAAGAACTTTTAAACACTACTACCTAGA 543
QY 228 GluGlnIleGluSer---AsnIleValArgCysProAspProGluTyrAlaGluLysMet 246
DB 544 GAACAAGTCGATCGTGTAGTCCCAATTAAGATGTCACAGATGCAAAATGTCGCGTGAAGAATG 603
QY 247 IleAspAlaIleAspArgValArgValArgLysAspSerValGlyValIleThrCys 266
DB 604 GTTAAAGTATTGAAATAATCTGTATGCTGTAAGATTCATTTGGTGTGTGCTATTGT 663
QY 267 ValAlaArgAsnValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGlu 286
DB 664 GTTATCAGAAATTTGTCATTTGATTTAGGTGAACCATGTTTCGATATAATTTGGAAGCTAAA 723
QY 287 LeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPhe 306
DB 724 TTGGCTCATGCCATGTTATCATTTACCCGCTACCAAGGGGTTGAATTTGGTTCGTGATTT 783
QY 307 AlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPheTyrMetAspLys---Ala 325
DB 784 GAAGGTATCAAGATCCTGTTCAAAACATAACATGATGTCATTTATGATGAAATTC 843
QY 326 GlySerValArgThrArgThrAsnArgSerGlyValGlnGlyIleSerAsnVal 345
DB 844 CGAAGATTAAAGACTGAACAAACAATAGTGTGTGATTCAGAGGAGATCTCAACGCT 903
QY 346 GluIleValHisPheLysValAlaPheLysProThrProSerIleGlyValLysGlnAsn 365
DB 904 GAAACATTTATTCTCAGTTGCTTCAATTCGAGTGCCTACTACTAGTCAAGAACCAAGAA 963
QY 366 ThrValSerArgGluArgGlnAsnValGluLeuAlaArgGlyValArgHisAspProCys 385
DB 964 ACTGCCCATTTATGTTGTTAAAGTGTGTTTGGCTGTAGAGTAGACACCATCAAGT 1023
QY 386 ValAlaProArgAlaValProValValGluSerMetAlaAlaLeuValLeuMetAspGln 405
DB 1024 GTCAACCCCAAGAGCAGTTCCAATTTGCGAAGCCATGCTACTGTTATGTTGAGCA 1083
QY 406 LeuMetAlaHisValAlaGln 412
DB 1084 TATATGATTCAACAAAGCTAGA 1104
```

RESULT 15

```
ADR85489
ID ADR85489 standard; DNA; 1239 BP.
XX
AC ADR85489;
XX
DT 04-NOV-2004 (first entry)
XX
DE Aspergillus fumigatus essential gene open reading frame #126.
XX
KW Fungicide; Aspergillus fumigatus infection; Farmer's lung disease;
XX drug screening; ds.
XX
OS Aspergillus fumigatus.
XX
PN WO2004067709-A2.
XX
PD 12-AUG-2004.
XX
PF 16-JAN-2004; 2004WO-US001099.
XX
PR 17-JAN-2003; 2003US-0441281P.
XX
PR 13-JUN-2003; 2003US-0478196P.
XX
PA (ELIT-) ELITRA PHARM INC.
PA (ELIT-) ELITRA CANADA LTD.
XX
PI Jiang B, Hu W, Lemieux S, Roemer T;
XX
WPI; 2004-S94200/57.
P-PSDB; ADR86076.
XX
PT New purified or isolated Aspergillus fumigatus nucleic acid molecule
PT encoding a gene product, useful for diagnosing and/or treating invasive
PT fungal infections, such as Farmer's lung disease.
XX
PS Claim 2; SEQ ID NO 2126; 164pp; English.
XX
CC The present invention relates to Aspergillus fumigatus genes that are
CC essential and are potential targets for drug screening. The methods and
CC compositions of the present invention are useful for diagnosing and/or
CC treating invasive Aspergillus fumigatus infection, including the allergic
CC forms of the disease, such as Farmer's lung disease. They can also be
CC used in various drug discovery purposes, such as expression of the
CC recombinant protein, hybridization assay and construction of nucleic acid
CC arrays. The present sequence represents an Aspergillus fumigatus
CC essential gene open reading frame, used during diagnosis and drug
CC development in the invention. These genes share a high degree of sequence
CC conservation with known essential genes of candida albicans. The sequence
CC data for this patent is not represented in the printed specification, but
CC was obtained in electronic format from WIPO.
XX
SQ Sequence 1239 BP; 255 A; 431 C; 330 G; 223 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6 9e-76 Length: 1239
Score: 929.50 Matches: 196
Percent Similarity: 66.49% Conservative: 58
Best Local Similarity: 51.31% Mismatches: 105
Query Match: 42.25% Indels: 23
DB: 13 Gaps: 6

US-10-677-179-8 (1-429) x ADR85489 (1-1239)
QY 54 AnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGlyGlyVal 73
DB 4 TCGACGTGGGAGAACTATCTTCGGGTCCACCATGTGTGAATCCATTCGCGCTCTGTC 63
QY 74 GlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAspLeuGlnVal 93
DB 64 GGTCTGATCGTGTGATGCTGCCCTCCAGGCATGGAGCTTACAGAGAAAGACATCCAAACC 123
QY 94 GluLeuAspArgArgArgProGlyGlnSerArgIleThrSerThrArgLysGluThrAsp 113
```

Job time : 639 secs

```
Db      124  CAGATGACTCGAAGACGTCCTCCGGGAGAGTGTCTTAACGACGCCTCGAAATGAAAAAGAC 183
QY      114  ThrCysLysIleLeuSerGlyThrHisGluGlyValThrGlyThrProIleLeuVal 133
Db      184  CGAGTAGAGATCCAGTCTGGAACGAGTTCCGTCATCACCTGGGTACCCCGATTGGAATG 243
QY      134  IleValProAsnThrAspGlnIleGlySerAspHis---ArgGluIleAlaAsnValTyr 152
Db      244  ATGGTGCACACGAGGATCAGAGACCAAGGACTACGGTGGCGAGCACCAATGGATCTCTAC 303
QY      153  ---ArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGln 171
Db      304  CCTCGTCCCGTCAGCTGATTACTTACCTGGAGAAATACGGTGTCAAGCGGAGCAGC 363
QY      172  GlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuPro 191
Db      364  GGTGGTGGCCGGAGTAGTCCCGCGAGACCATGGCGGTGTCGCCGAGGAGCCATTGGC 423
QY      192  LysLysIleLeuLysCysGlyLeuGluIleLeuSerPheValSerLysValHis 211
Db      424  GAGAAGTACCTACGCTGTCGTCATGGTGTGAAATTTGTGCGCTTTGTCTCGCTGGT 483
QY      212  GlnValValLeuProGluAspAlaValAspTyrGlySerValThr----- 226
Db      484  AAGAACACCTTTTCCGCGGACCCCGAGCACCTTCTCCATCGACCAACCCCTGAGTTC 543
QY      227  -----LeuGluGlnIleGluSerAsnIleVal-----ArgCys 237
Db      544  CTGAAGCTCATCGAGACCATCGACCGTAAGACTGTCGATGCTTCGTCCCGACTCGCTGC 603
QY      238  ProAspProGluTyrAlaGluLysMetIleAspAlaIleAspArgValArgValArgGly 257
Db      604  CCGAACGAGGAGCGCGCGCATGACAAAGTGCATCGAGACTTTCGGGACCAACCAA 663
QY      258  AspSerValGlyGlyValIleThrCysValAlaIleArgAsnValProArgGlyLeuGlySer 277
Db      664  GATAGCATCGCGGACCGTACCTCGGTGATCCGCAAGCTCCCGTGGCTGGCGCGAG 723
QY      278  ProValPheAspLysLeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSer 297
Db      724  CCTTGTTCGACAAGCTCGAGGCCAAGCTGGCGCACGCCATGCTCAGCATCCCGCCACC 783
QY      298  AsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsn 317
Db      784  AAGCGCTTTGAGATCGGCTCGGGCTTCGGTGGCTCGAGGTCCCGGGCTCCATCCACAA 843
QY      318  AspGluPheTyrMetAsp-----LysAlaGlySerValArg-----Thr 330
Db      844  GACCTCTTCCCGTCTCCGAGGTCCAGACCCGCGCAGCGCAGCACAGCGCTGACCACC 903
QY      331  ArgThrAsnArgSerGlyGlyValGlnGlyIleSerAsnValGluIleValHisPhe 350
Db      904  AAGACCAACAACCTCCGGCGGCATCCAGGCGGGATCTCCACGCGGCTCCCATCTATTTC 963
QY      351  LysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGlu 370
Db      964  CGCGTTGCCCTTCAAGCCCCCGCCACCATCGGCCAAGGCTCAGACCACCGCCTCTTACAG 1023
QY      371  ArgGlnAsnValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaProArgAla 390
Db      1024  TTCGAGGAGGGATCTCCGAGGCCAAGGCCCGCCAGACCCCTGGTTACCCCTCGTGCT 1083
QY      391  ValProValIleGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisVal 410
Db      1084  GTCCCCATCGTCGAGGCCATGTCCGCCCTCGTCTGTCATGATGCGCTCATGGCCCGCAGTAT 1143
QY      411  AlaGln 412
Db      1144  GCCCGC 1149
```

Search completed: August 27, 2005, 04:55:06

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 27, 2005, 04:20:24 ; Search time 3357 Seconds
(without alignments)
4864.335 Million cell updates/sec

Title: US-10-677-179-8

Perfect score: 2200

Sequence: 1 MTTVPKQQVAHSRARLAPR.....VAQCEMFALNTALQBPVGSF 429

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 15032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10677179/runat_25082005_150240_24847/app_query.fasta_1.583
-DB=EST -OPMT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10677179 @CN 1_1 5180 @runat_25082005_150240_24847 -NCFU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

```

1: gb_est1:*
2: gb_est2:*
3: gb_hc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2200	100.0	1626	3	AY105171 Zea mays
2	1441	65.5	1488	3	CNS0ACWT BX15746 Arabidops
3	1434	65.2	1494	3	CNS0ACWT BX18532 Arabidops
4	1398	63.5	1538	3	CNS0ACWT BX14847 Arabidops
5	1182.5	53.8	1160	7	CK210372 FGAS02217
6	1170	53.2	923	7	CO082566 GR_Ea47E
7	1135	51.6	763	6	CB654761 OSJNEC071
8	1120.5	50.9	911	7	CV256756 WS0244.B2
9	1087	49.4	865	7	CO114396 GR_Eb015

10	1042	47.4	786	4	BG600177
11	1039	47.2	835	4	BM411032
12	1032	46.9	710	4	BJ268020
13	1027.5	46.7	814	7	CNI44654 WJNDI_23
14	1022	46.5	723	7	CF303160
15	1022	46.5	755	7	CK279831 EBF1--01-
16	1018	46.3	835	7	CO080629 GR_Ea44D
17	1011	46.0	702	5	BU099249 WHE1305.C
18	996.5	45.3	919	7	CN385253 LE2TR03B1
19	990	45.0	869	7	CO933258 GR_Ea14L
20	989	45.0	864	6	CAL89902 SCCCLR1C0
21	988	44.9	803	6	CB066694 EST645317
22	985	44.8	690	6	CA124299 SCQGLR108
23	984	44.7	667	6	CAL19896 SCCCLR106
24	970	44.1	818	7	COL132399 GR_Eb45D
25	967	44.0	694	6	CD870292 AZO2_113P
26	959	43.6	812	7	CNI29049 RHOH1_33
27	954	43.4	775	7	CO101191 GR_EB002
28	948	43.1	696	5	BQ841547 WHE4212.D
29	946	43.0	783	7	COL12743 GR_Eb004
30	932	42.4	871	7	CO090892 GR_Ea11B
31	929	42.2	673	7	CF571696 MCS016H05
32	924	42.0	689	6	CA298301 SCRFPL803
33	921	41.9	726	6	CD008025 VVB082A11
34	918	41.7	639	4	BG414940 HVSMEX000
35	918	41.7	722	4	BI930866 EST550755
36	916.5	41.7	749	6	CA094750 SCCCL401
37	916	41.6	957	7	CV263720 WS02022.B
38	908	41.3	712	6	CD712988 VVB13A07
39	906	41.2	708	6	CD009988 VVB102E11
40	901	41.0	682	2	AW223745 EST300556
41	901	41.0	707	6	CD714440 VVB200H12
42	895	40.7	605	6	CA220612 SCRFPL400
43	893	40.6	907	7	CO074131 GR_Ea34B
44	892	40.5	661	6	CD899216 GI74_111J
45	891	40.5	629	1	AV832717 AV832717

ALIGNMENTS

RESULT 1	AY105171	AY105171	1626 bp	linear	HTC 16-OCT-2002
LOCUS	Zea mays	PCO087106	mRNA		
DEFINITION	Zea mays	PCO087106	mRNA sequence.		
ACCESSION	AY105171				
VERSION	AY105171.1	GI:21208249			
KEYWORDS	HTC.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.				
AUTHORS	Hayney,C.P., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.				
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes				
JOURNAL	Unpublished (2002)				
REFERENCE	2 (bases 1 to 1626)				
AUTHORS	Coe,E.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA				
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.				
FEATURES	Location/Qualifiers				
source	1..1626				
	/organism="Zea mays"				

```

/mol_type="mRNA"
/db_xref="MaizeDB:634903"
/db_xref="caxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN
Alignment Scores:
Pred. No.:      3,826-225      Length:      1626
Score:          2200.00      Matches:      429
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              3          Gaps:      0

US-10-677-179-8 (1-429) x AY105171 (1-1626)
QY 1 MetThrValProLysProGlnGlnValAlaHisSerArgAlaArgLeuAlaProArg 20
DB 79 ATGACGCGTCCCAAGCCACAGCAGGTGGGCACTCAGGGCAGCGCTCGACCCCGC 138
QY 21 AlalLeGlyAlaLeuLeuGluPheAlaProLaserSerLeuArgPheAlaValHis 40
DB 139 GCGATCGCGCGCTTCTGCGAGTTTCCGCCAGCCTCCTCCTCCGCTTCCGCGTGCA 198
QY 41 ArgCysArgThrAlaArgLeuGluValLysAlaSerGlyAsnThrPheGlyAsnThrPhe 60
DB 199 CGCTCCCGCAGCTGCTGGCTAGAGTGAAGGCATCTGGAAACACGTTTGGAAACTACTTT 258
QY 61 GlnValAlaThrTyrGlySerHisGlyGlyValGlyCysValIleSerGlyCys 80
DB 259 CAGGTTCACACCTATGGTGAATCTCATCGGGGTGGTGTGGTGTATCAGTGGTGT 318
QY 81 ProProArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArgArgPro 100
DB 319 CCACCTAGAAATTCACCTCACTGAGGCAGACCTTACAAAGTTGAACCTCGATCGAAGACGGCCC 378
QY 101 GlyGlnSerArgIleThrSerThrArgLysGluThrAspThrCysLeuIleLeuSerGly 120
DB 379 GGACAGAGCAGAAATAACTCCCAAGGAGGAGACTGTATACATGCAAAATTTCTGTGAGGG 438
QY 121 ThrHisGluGlyValThrThrGlyThrProIleLeuValIleValProAsnThrAspGln 140
DB 439 ACACATGAAGGGGTGACTACTGGAAGGCCAATCTTGTATTGTGCCAAACACAGATCAA 498
QY 141 IleGlySerAspHisArgGluIleAlaValThrArgProSerHisAlaAspAlaThr 160
DB 499 ATAGGCAGTGATCACCGTGAATAGCAATGTGTACCGACCTTCTCATGCGACAGCAACT 558
QY 161 TyrAspPheLysTyrGlyValArgAlaValGlnGlyGlyArgSerSerGlyArgLys 180
DB 559 TATGACTTCAAGTACGGTGTATAGAGCTGTACAGGAGGTTGGAGGTCTCTCGGGCGAANA 618
QY 181 ThrValGlyArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysLeuLysCysGly 200
DB 619 ACCGTTGGAAGGGTGGCTGCAGGGGCCCTCCCAAGAAATTTCTTAAGCTCAATGTGGA 678
QY 201 LeuGluLeuLeuSerPheValSerLysValHisGlnValValLeuProGluAspAlaVal 220
DB 679 TTAGAGATCTTGTGCTTTTCCAAAGTGCATCAGGTGTGTCTCCCGAAGACGCGTT 738
QY 221 AspTyrGlySerValThrLeuGluGlnIleGluSerAsnIleValArgCysProAspPro 240
DB 739 GATTATGGGTCTGTAACTTTGGAACAGATAGAGCAACATCGTTAGATGTCTCTATCCA 798
QY 241 GluTyrAlaGluLysMetIleAspAlaIleAspArgValArgValArgGlyAspSerVal 260

```

```

Db 799 GAGTACGACAGAGAAGATGATAGACGCAATCGACAGAGTACGAGTTCGAGGGGATTCGGTC 858
QY 261 GlyGlyValIleThrCysValAlaAlaArgAsnValProArgGlyLeuGlySerProValPhe 280
DB 859 GGTGGAGTGTATCATCGCTCGCTAGAAACGTTCTCTCGCGGGTCTCGTTCCTGTGTTTC 918
QY 281 AspLysLeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPhe 300
DB 919 GACAGCTCGAATCCGAACTGGCAAAAGCTATGTTCTTATTCTCTGGCAGCAACGGGTTTC 978
QY 301 GluIleGlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPhe 320
DB 979 GAGATTGCGACGCGGATTTCGCGCGGACCGACTTCACAGCAAGTGTAGCATAATGATGACTTT 1038
QY 321 TyrMetAspLysAlaGlySerValArgThrArgThrAsnArgSerGlyGlyValGlnGly 340
DB 1039 TATATGGATAAGGCTGGAACTGTGAGGACACGCACTAATCGCTCGGGTGTGTGCAGGGA 1098
QY 341 GlyIleSerAsnValGluIleValHisPheLysValAlaPheLysProThrProSerIle 360
DB 1099 GGGATATCGAATGTTGAGATTGTGCACTTCAAGTTGCTTTTAAGCCGACACCATCTATC 1158
QY 361 GlyValLysGlnAsnThrValSerArgGluArgGlnAsnValGluLeuLeuAlaArgGly 380
DB 1159 GGGGTGAACAGAAACACCGTGTCAAGGGAGCGTCAGAACTGTGAGCTTCTAGCAAGAGGG 1218
QY 381 ArgHisAspProCysValAlaProArgAlaValProValValGluSerMetAlaAlaLeu 400
DB 1219 CGCATATGACCATCGCTCGCCCTCGAGCTGTCTGTGTGGTAATCCATGGCGCGCTTG 1278
QY 401 ValLeuMetAspGlnLeuMetAlaHisValAlaGlnCysGluMetPheAlaLeuAsnThr 420
DB 1279 GTCTCTATGACCACTGATGGCGCAGTGGCTCAGTGCAGAGATGTTTCGGCTCAACT 1338
QY 421 AlaLeuGlnGluProValGlySerPhe 429
DB 1339 GCATTCAGAACACAGTTGGCTCTTTC 1365

RESULT 2
CNSOACWT 1488 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSUTLS82607 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION BX815746
VERSION BX815746.1 GI:42473095
KEYWORDS HTC; GSLT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1488)
AUTHORS Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1488)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
REFERENCE BP 191006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr)
AUTHORS - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.

```


Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Fulllength

<http://www.genoscope.cns.fr/cgi-bin/qgb/qgb?source=Arabidopsis>.

```

FEATURES
source
1. .1488
Location/Qualifiers
1. 1488

```

Source

Location,
1 1400

1. 1488

```

/organism="Arabidopsis thaliana"

```

/mol type="mRNA"

/strain="Col-0"

/db_xref="taxon:3702"

```
/clone="GSLTLS8ZG07"
```

/tissue_type="Adult vegetative tissue"

```
/plasmid="pCMVSPORT_6"
```

complement(1. .1488)

Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length
<http://www.genoscope.cns.fr/cgi-bin/qgb/qgb?source=Arabidopsis>.

FEATURES

```
1. .1494
/mol_name="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/cdone="GSLRS1842805"
/tissue_type="Silique"
/plasmid="pCMWSPORT 6"
complement(1. .1494)
/gene="At1g48850"
```

ORIGIN

Alignment Scores:				
Pred. No.:	7,25e-143	Length:	1494	
Score:	1434.00	Matches:	267	
Percent Similarity:	83.63%	Conservative:	65	
Best Local Similarity:	67.25%	Mismatches:	64	
Query Match:	65.18%	Indels:	1	
DB:	3	Gaps:	0	
US-10-677-179-8 (1-429) x CNSOACNM (1-1494)				
Qy	31	AlaSerSerSerLeuArgPheAlaValHisArgCysArgThrAlaArgLeuGluVallys	50	
Db	163	TCCTCTCCCGCGTTCAGATCTCTCTCGTACCCAAACCAGGAGAACTTCCAGATACAA	222	
Qy	51	AlaSerGlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGly	70	
Db	223	GCTACTGGGAAGTTCATATGGGACTCATTTTCGAGTTTCAACTTTGGAGAATCACATGGA	282	
Qy	71	GlyGlyValGlyCysValIleSerGlyCysProProArgIleProLeuThrGluAlaAsp	90	
Db	283	GGAGGAGTTGGTTGTATCATTTGATGGTTGCTCTCGTATTCACCTACTCGAATCTGAT	342	
Qy	91	LeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLys	110	
Db	343	TTTCAATTCCGATCTCGATAGAGGAGGCTGGTCAGAGCAGGATCACAACTCCTAGAAAA	402	
Qy	111	GluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThrPro	130	
Db	403	GAGACTGATACTGCCGATATCGTCTGGAGTCTCTGAGGAATGACGACAGGAACACCT	462	
Qy	131	IleLeuValIleValProhenThrAspGlnIleGlySerAspHisArgGluIleAlaAsn	150	
Db	463	ATCCATGTGTTGTATACCAACACATCAGAGAGGACTTGATTACAGTGAATGTCGGTT	522	
Qy	151	ValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaVal	170	
Db	523	GCCTATAGACCATCGCATGCTGATCCAACTTATGACATGAATGATGTTGCATCAGTG	582	
Qy	171	GlnGlyGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeu	190	
Db	583	CAGGGTGGAGGAAGATCTTTCAGTAGAGAGACCATTGGAGAGATGTGCTCCTCGAGCTTG	642	
Qy	191	ProLysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysVal	210	
Db	643	GCCAGAGAAATTTTGAGCCATTTCAGGAACCTAGATTTCTGCTATGTCTCGCAATG	702	
Qy	211	HisGlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIle	230	
Db	703	CACCATGTGTATCTTCCAGAGAAATGTGTAGACACACAGAAATTTAACTCGAACAGAT	762	
Qy	231	GluSerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIle	250	
Db	763	GAATAATACATTTGTAGATGCCCTATATCCGAGTATGCCGAAAGATGATAGCTGCGATT	822	
Qy	251	AspArgValArgValArgValAspSerValGlyGlyValIleThrCysValAlaArgAsn	270	

```

length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
FEATURES
    source                1..1538
                        /organism="Arabidopsis thaliana"
                        /mol_type="mRNA"
                        /strain="Col-0"
                        /db_xref="taxon:3702"
                        /clone="GSLTSL102H06"
                        /tissue_type="Adult vegetative tissue"
                        /plasmid="pQWSPORT 6"
                        complement(1..1538)
                        /gene="Atlg48850"
ORIGIN
Alignment Scores:
Pred. No.:      5,59e-139      Length:      1538
Score:          1398.00      Matches:      267
Percent Similarity: 83.17%      Conservatives: 64
Best Local Similarity: 67.09%      Mismatches: 65
Query Match:      63.55%      Indels: 3
DB:                3          Gaps: 0
US-10-677-179-8 (1-429) x CNS0ACYT (1-1538)
QY  31  AlaSerSerSerLeuArgPheAlaValAlaHisArgCysArgThrAlaAArgLeuGluValLys 50
Db  222 TCCTCTCCCGCGGTTCAGATCTCTCTCCGTACCCAAACAGGAAGAACTTCCAGATACAA 281
QY  51  AlaSerGlyAsnThrPheGlyAsnThrPheGlnValAlaThrTVrGlyGluSerHisGly 70
Db  282 GCTACTGGGAAGTTCATATGGGACTCATTTTCAGTTTCACCTTTTGGAGAAATCACATGGA 341
QY  71  GlyGlyValGlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAsp 90
Db  342 GGAGGAGTTCGTTGTATCATTCATTCATGCTGTCTCTCTCGTATTCACCTTACTGAAATCTGAT 401
QY  91  LeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLys 110
Db  402 TTGCAATTGCATCTCGATGAAGGAGGCTCGTCAGACGAGGATCACAACTCTCTGAGAAA 461
QY  111 GluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThrPro 130
Db  462 GAGACTGATACTTCGCGGATATCGTCTCGAGTCACTGAAGGAATGACACAGGAACACT 521
QY  131 IleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsn 150
Db  522 ATCCATGTGTTGTATACCAACACACATCAGAGAGGACTTGATTACAGTGAATATGCGGTT 581
QY  151 ValTVrAtqProSerHisAlaAspAlaThrTVrAspPheLysTVrGlyValArgAlaVal 170
Db  582 GCCTATAGACCATCGATGCTGATGCCAACTTATGACATGAAGTANGTGTGCAGTCAGTG 641
QY  171 GlnGlyGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeu 190
Db  642 CAAGGTGGAGGAAGATCTTCAGCTAGACAGACCACTTCGGAGAGTTCCTCTCGGAGCTTGT 701
QY  191 ProLysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysVal 210
Db  702 GCCAAGAAATT-TTGAAGCAATTTTCAGGAACTCAGATTTCTCGCTATGTCTCGCAAGTC 760
QY  211 HisGln-ValValLeuProGluAspAlaValAspTVrGlySerValThrLeuGluGlnI 230
Db  761 CACCATTTGTTACTCCAGAGAAATTTGGTAGACCACAGAAATTTTACACTTCGACAGAT 820
QY  230 eGluSerAsnIleValArgCysProAspProGluTVrAlaGluLysMetIleAspAlaI 250
Db  821 AGAAATAACATTTGTTCAGATCCCTTAATCCGAGTATCGGAAAGAAATGATAGTCCGAT 880
QY  250 eAspArgValAlaArgGlyAspSerValGlyValIleThrCysValAlaArgAsn 270
Db  881 TGATCTGTGAGGACAAAGGAAACTGTGTGGTGGTGTGTGACCTGCATGTTGTCGAA 940

```

of spike formation in wheat cultivar Norstar. 4 mRNA populations were combined before constructing the library. The first mRNA population is from 1cm crown sections after 30 days of cold acclimation. The second is from 1cm crown sections after 11 days of deacclimation (before deacclimation plants were fully vernalized for 49 days). The third is from different developmental stages of spike formation (5 to 50mm) that still have not emerged from the leaf (dissection required). The last is from different developmental stages of spike and seed formation after having emerged from the leaf (visible). First strand synthesis in this library was done in the presence of methylated dCTP thereby protecting from internal cleavage with NotI."

ORIGIN

Alignment Scores:

Pred. No.: 5,41e-116 Length: 1160
Score: 1182.50 Matches: 245
Percent Similarity: 76.92% Conservative: 35
Best Local Similarity: 67.31% Mismatches: 53
Query Match: 53.75% Indels: 32
DB: 7 Gaps: 5

US-10-677-179-8 (1-429) x CK210372 (1-1160)

```

Qy 2 ThrThrValProLysProGlnGln----- 9
Db 23 ACGGTCGCGCGAACCACCAACCAATCTACACCCCTCTCCAAATGGCAACGCGCCC 82
Qy 10 ValAlaHisSerArgAlaArgLeuAlaPro-----ArgAlaIleGly 23
Db 83 ACGTGCGACACAGTCGCGCGCGCGCGCCCTGGAGCTCCCTCCCGCGGGGTTCCGG 142
Qy 24 AlaLeuLeuGluPheAlaProAlaSerSerSerLeuArgPheAlaValHisArgCysArg 43
Db 143 GCGTCTCAGGACTCGCGCGCGCGCC-----TCCGTCCGTCTTCCGTCCGTCCGCGCGC 196
Qy 44 ThrAlaArgLeuGluValLysAlaSerGlyAsnThrPheGlyAsnThrPheGlnValAla 63
Db 197 GCCCGCGCGCTAGAGGTGAAGACGCTGGAAACATCTTCGGGGACTACTTCCAGGTGCA 256
Qy 64 ThrYrGlyGluSerHisGlyGlyValGlyCysValLysSerGlyCysProProArg 83
Db 257 ACTTATGAGAGTCTCATGGAGCGGTGTGGCTGCGTTATCATGTGTGTCACCCAGA 316
Qy 84 IleProLeuThrGluAlaAspLeuGlnValGluLeuAspArgArgArgProGlyGlnSer 103
Db 317 ATCCCACTCAGGAGGAAGACATG-----CAGAGC 346
Qy 104 ArgIleThrSerThrArgLysGluThrAspThrCysLysLysLysLeuSerGlyThrHisGlu 123
Db 347 AGAATAACACCCCAAGGAGGAGGACGATACTTGTAAATCTCTTCAGGACATACGAA 406
Qy 124 GlyValThrThrGlyThrProIleLeuValLysValProAsnThrAspGlnIleGlySer 143
Db 407 GGAATGACCACTGGGACACCAATTCATGTTCTTGTCCCAACACGATCAAGAGGGGT 466
Qy 144 AspHisArgGluIleAlaAsnValTyrArgProSerHisAlaAspAlaThrTyrAspPhe 163
Db 467 GATTACACTGAAATGGCTAAGCGGTACAGACCTTCCCATCGGATTTAACTTATGACCTC 526
Qy 164 LysThrGlyValArgAlaValGlnGlyGlyArgSerSerGlyArgLysThrValGly 183
Db 527 AAATACCGGTGTAGATCTGTGACGGAGGTGGAAAGGTCATCGGCAAGAGAAACCATTTGA 586
Qy 184 ArgValAlaAlaGlyAlaLeuProLysLysLysLysLysLysLysGlyLeuGluIle 203
Db 587 AGGTAGCTGACGAGCTGTGTGCAAGAAATTTTAAGCTCAATGTGGGGTAGAATT 646
Qy 204 LeuSerPheValSerLysValHisGlnValValLeuProGluAspAlaValAspTyrGly 223
Db 647 CTAGCAATTTGTTTCCAAAGTGCATCAAGTGTACTTCTTGAAGACGCGAGTTGATTACGAA 706

```

```

Qy 224 SerValThrLeuGluGlnIleGluSerAsnIleValArgCysProAspProGluTyrAla 243
Db 707 GCTCTTACCCTGTGATCAGATAGAGAACATCTGTAGATGTCTGATCCAGAAATATGCA 766
Qy 244 GluLysMetIleAspAlaIleAspArgValArgValArgGlyAspSerValGlyVal 263
Db 767 CAGAAGATGATTCATGCAATGATAAAGTACGAATTAATGGGAATTCGATTGTTGGGTG 826
Qy 264 IleThrCysValAlaArgAsnValProArgGlyLeuGlySerProValPheAspLysLeu 283
Db 827 GTCACATGCAATGCCAGAAATGCTCTCGTGGGCTTGTCTCTGTATTGCAAACTG 886
Qy 284 GluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPheGluIleGly 303
Db 887 TGAGCTCTACTGCAAGGCTATGCTTTCTTCTTCCGCAAGCAAGGGGTGAGATCCGT 946
Qy 304 SerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPheTyrMetAsp 323
Db 947 CATGGGAATTCAGGTACTGACCTCTACTGGAGTTGAGAT-AACGATAAGTTCTATATGGAT 1005
Qy 324 LysAlaGlySerValArgThrArgThrAsnArgSerGlyGlyValGlnGlyIleSer 343
Db 1006 TAGGTTGAAATGCTGAAACACCAACCAATCATG---CTCGCGGGGTTCCAGGAGGATATCA 1062
Qy 344 AsnValGluIle 347
Db 1063 AATGGGGAACATA 1074

RESULT 6
LOCUS CO082566
DEFINITION GR_Ea47E01.r GR_Ea Gossypium raimondii cDNA clone GR_Ea47E01 3', mRNA sequence.
ACCESSION CO082566
VERSION CO082566.1 GI:48752047
KEYWORDS EST.
SOURCE Gossypium raimondii
ORGANISM Gossypium raimondii
REFERENCE 1 (bases 1 to 923)
AUTHORS Kim H., Yu Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C., Udall, J.A., Rapp, R.A., Wendel, J.F., Rao, K., Soderlund, C. and Wing, R.A.
TITLE Global assembly of Cotton ESTs
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
Plate: 47 row: E column: 01.
FEATURES
source
1..923
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Ea47E01"
/tissue_type="whole seedlings"
/dev_stage="first true leaves"
/lab_host="DH10B"
/clone_lib="GR_Ea"
/notes="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2: EcoRV; Library made by Invitrogen with RNA supplied by Wendle lab. Directional cloned into NotI-EV. Colonies plated/picked by AGI. More glycerol clones held in -80."
ORIGIN
Alignment Scores:

```

Pred. No.: 8.6e-115 Length: 923
 Score: 1170.00 Matches: 221
 Percent Similarity: 87.30% Conservative: 47
 Best Local Similarity: 71.9% Mismatches: 39
 Query Match: 53.18% Indels: 1
 DB: 7 Gaps: 0

US-10-677-179-8 (1-429) x C0082566 (1-923)

QY 112 ThrAspThrCysLeuValLeuSerGlyThrHisGluGlyValThrGlyThrProIleu 131
 DB 2 ACAGACACATGCCGAATATATCTGGACTTCTGAAGGAGTGACTACTGGACACCAATC 61
 QY 132 LeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnVal 151
 DB 62 CATGTACTGTACCAATACTGATCAGAGAGACATGATTACAGGAATGTCAATAGCT 121
 QY 152 TyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGln 171
 DB 122 TATAGGCCATCTCATGCTGATGCCACTTATGACATGAAATATGGTGTGAGGCGAGTGCAG 181
 QY 172 GlyGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuPro 191
 DB 182 GGTGTGTGTAGATCATCGCCAGAGAAACCATTTGGAGAGTTGCTCTGGAGCTATTGCT 241
 QY 192 LysLysIleLeuLysLysCysGlyLeuGluIleLeuSerPheValSerLysValHis 211
 DB 242 AAGAAATCTCAAGCTGTTTCAGAACTGAGGTCTTGCTATGTCTCTCAAGTTTCAAC 301
 QY 212 GlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGlu 231
 DB 302 CAGGTGTGTTTCCAGATGGCTCAGTGCACATGACACTGTAACTCTTGTATCAGATAGAG 361
 QY 232 SerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIleAsp 251
 DB 362 AGCAATATGTGAGGTGCCAAATCTGATTACGCTGAGAAATGATTGCTGCTATTGAT 421
 QY 252 ArgValArgValArgGlyAspSerValGlyValIleThrCysValAlaAlaArgAsnVal 271
 DB 422 GCTGTGCGGACNAGAGAAATCTATTGCGGAGTTGTACATGATAGTAGGAGATGCT 481
 QY 272 ProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMet 291
 DB 482 CCAGTGGCTTGGTTTACCAGTTTTCATAGCTTTGAAGCGGAGCTTGTCTAAGGCTGTA 541
 QY 292 LeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeu 311
 DB 542 ATGTCACTACTGCAACCAAGGCTTTGAATTTGGGAGTGGATTTGCAGGTACTTTGTTG 601
 QY 312 ThrGlySerGluHisAspGluPheTyrMetAspLysAlaGlySerValArgThrArg 331
 DB 602 ACCGGTAGTGAACATATATGATGAGTTCTATCTGATGATCACTGAGGAATAGGACAGA 661
 QY 332 ThrAsnArgSerGlyValGlnGlyGlyLysSerAsnValGluIleValHisPheLys 351
 DB 662 ACNAATGCTCCGGTGGGATACAGGCTGGAATATCCAAATGGGAAATATTAATATGAGA 721
 QY 352 ValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGluArg 371
 DB 722 GTAGCTTTCAAGCCCAACAGCTACAAATTTGGTGAAGAAACAGCACACAGTGACTCGAGAAA 781
 QY 372 GlnAsnValGluLeuAlaArgGlyArgHisAspProCysValAlaAlaProArgAlaVal 391
 DB 782 GAAGNATAGAACTACTAGCCGCGGTGCTCATGATCTCTTGTGTCTCCCTCGAGCGGTG 841
 QY 392 ProValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisValAla 411
 DB 842 CCGATGGTTGAAGCGATGTAGCGTGTGCTGTGCTGTGTTGTTGCTCCCTCGAGCGGTG 901
 QY 412 GlnCysGluMetPheAlaLeu 418
 DB 901 CAATGTAACTTTTCCCTTATC 921

RESULT 7
 CB654761
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

CB654761 763 bp mRNA linear EST 09-APR-2003
 OSJNEC07113.f OSJNEC Oryza sativa (japonica cultivar-group) cDNA
 clone OSJNEC07113 5', mRNA sequence.

CB654761
 EST
 GI:29658486

Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.

1 (bases 1 to 763)

Jantasuriyarat, C., Lu, G., Gowda, M., Hatfield, J., Zhou, B., Mazur, E.,
 Kudrna, D., Dean, R., Soderlund, C., Wing, R. and Wang, G.

Large-scale identification of ESTs involved in the interaction
 between rice and Magnaporthe grisea

Unpublished (2003)

Contact: Rod Wing

Arizona Genomics Institute

University of Arizona

Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ

85721-0088, USA

Tel: 520 626 3967

Fax: 520 621 9288

Email: http://genome.arizona.edu

PCR Primers

FORWARD: gta aaa cga cgg cca gtg

BACKWARD: gga aac agc tat gac cat g

Plate: 07 row: I column: 13

Seq primer: gta aaa cga cgg cca gtg.

Location/Qualifiers

1..763

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clone="OSJNEC07113"

/tissue_type="Leaf"

/dev_stage="3 week"

/lab_host="DH10B"

/clone_lib="OSJNEC"

/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:

XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"

ORIGIN

Alignment Scores:

Pred. No.: 3.78e-111 Length: 763
 Score: 1135.00 Matches: 214
 Percent Similarity: 93.70% Conservative: 24
 Best Local Similarity: 84.25% Mismatches: 16
 Query Match: 51.59% Indels: 0
 DB: 6 Gaps: 0

US-10-677-179-8 (1-429) x CB654761 (1-763)

QY 113 AspThrCysLysIleLeuSerGlyThrHisGluGlyValThrGlyThrProIleLeu 132
 DB 1 GACACTTGCAAAATTCCTTTCAGGACACATGAGGAATGACCACTGGGACACCAATTCAT 60
 QY 133 ValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnValTyr 152
 DB 61 GTTTTGTCCGGAACACAGATCAGAGAGGGGTGATTACAGTGAATGGCTTAAGGCCTAC 120
 QY 153 ArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGlnGly 172
 DB 121 AGACCTTCACATGCAGATGCAACTTATGACTTCAATACGGTGTTAGAGCAGTGCAGGA 180
 QY 173 GlyGlyArgSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuProLys 192
 DB 181 GGTGGAAGATCATCAGCAAGAGAGACCATTTGGAAGGGTGGCTGCAGGAGCTCTTTCGAAAG 240

```

Qy 193 LysileLeuLysCysGlyLeuGluLeuSerPheValSerLysValHisGln 212
Db 241 AAAATTCCTTAAGCTCAAACTGGAGTAGAGATCTGGCGTTTGTCGAAGTGCATCAA 300
Qy 213 ValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGluSer 232
Db 301 GTTGACTACAGAGATGCCGTTGATATGACACTGTAAACATGGAACAGATAGAAAGC 360
Qy 233 AsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIleAspArg 252
Db 361 AACATTGTTAGATGCTCTGATCCAGATATGCACAGAGATGATGATGCAATCGATAAA 420
Qy 253 ValArgValArgGlyAspSerValGlyValIleThrCysValAlaArgAsnValPro 272
Db 421 GTACAGATTAGAGGTGATTCGATTGGTGTGTGTCACATGATTCGAAGAAATGTTCCCT 480
Qy 273 ArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMetIleu 292
Db 481 CGTGGATTGCTCTCTGTTATTTGACAACTTGAGGCTGAATGGCGAAAGCTATGCTT 540
Qy 293 SerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeuThr 312
Db 541 TCTCTTCCTGCAAGCAAGGGTTTGAGATCGGCACTGGAATTTGCAGGTACTGACTACACT 600
Qy 313 GlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThrArgThr 332
Db 601 GGAAGTGACATATGATGAGTTCATATGATGAGCTGGAATGTGGAACCAAGAACT 660
Qy 333 AsnArgSerGlyValGlnGlyIleSerAsnValGluIleValHisPheLysVal 352
Db 661 AATGTTTCAGCGGTGTGTCAGGAGGATATCAAAATGTTGAATATATATACTTTCAGAGTA 720
Qy 353 AlaPheLysProThrProSerIleGlyValLysGlnAsnThr 366
Db 721 GCTTTCAAGCCAACCGCACTATTGGGAAGAAACAACATACT 762

RESULT 8
CV256756/c
LOCUS CV256756 911 bp mRNA linear EST 22-SEP-2004
DEFINITION WS0244.B21_L13 PfxD-ICC-N-A-14 Populus balsamifera subsp.
trichocarpa x Populus deltoides cDNA clone WS0244_L13 3', mRNA
sequence.
ACCESSION CV256756
VERSION CV256756.1 GI:52509731
KEYWORDS EST.
SOURCE Populus balsamifera subsp. trichocarpa x Populus deltoides
ORGANISM Populus balsamifera subsp. trichocarpa x Populus deltoides
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE 1 (bases 1 to 911)
AUTHORS Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y.,
Kirkpatrick,R., Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G.,
Babakaliff,R., Brown-John,M., Chand,S., Featherstone,R., Maeson,A.,
Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A.,
Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and
Bohlmann,J.
The poplar transcriptome: Analysis of expressed sequence tags from
multiple cDNA libraries
JOURNAL Unpublished (2004)
COMMENT Contact: Joerg Bohlmann
Genome BC forest genomics program
University of British Columbia
UBC Biotechnology Laboratory, 6174 University Boulevard, Rm. 237,
Vancouver, British Columbia, Canada, V6T 1Z3
Tel: 1-604-822-0282
Fax: 1-604-822-6097
Email: bohlmann@interchange.ubc.ca
Plate: WS0244 row: L column: 13
High quality sequence stop: 911
POLYA=Yes.
FEATURES
Location/Qualifiers
source 1..911

```

/organism="Populus balsamifera subsp. trichocarpa x
Populus deltoides"
/mol_type="mRNA"
/cultivar="Hil-11"
/db_xref="taxon:3695"
/clone="WS0244_L13"
/sex="Male"
/lab_host="E. coli DH10B T1 phage resistant cells"
/clone_lib="PTXD-ICC-N-A-14"
/notes="Vector: pBluescript II SK (+) XR; Site 1: EcoRI (5'
end of cDNA); Site 2: XhoI (3' end of cDNA); Cultured
cells [de Sa MM et al. (1992) Plant Physiology 98:728-737]
were grown in media (45mM) supplemented with either 50uM
salicylic acid, 50uM benzothiadiazole, 50uM methyl
jasmonate, 20ug chitosan or 200uL of Pollacia radiosa
extract. Cells were harvested after a 3 hour treatment,
along with untreated control cells. mRNA was isolated from
each tissue source independently and equal quantities of
mRNA from each tissue were then pooled. cDNA was prepared
from 5 micrograms of mRNA and directionally ligated into
the pBluescript II SK (+) XR vector using the pBluescript
II XR cDNA Library Construction Kit according to
manufacturer's instructions with modifications
(Stratagene). Plasmid DNA was then transformed by
electroporation into DH10B cells (Invitrogen) for
propagation. Normalization was applied according to
published methods [Ronald M.P. et al. (1996) Genome
Research 6(9):791] in order to reduce the abundance of
highly expressed transcripts."

Alignment Scores: 1.76e-109 Length: 911
Pred. No.: 1120.50 Matches: 215
Score: 78.83% Conservative: 42
Percent Similarity: 65.95% Mismatches: 39
Best Local Similarity: 50.93% Indels: 30
Query Match: 7 Gaps: 1
DB: 1

US-10-677-179-8 (1-429) x CV256756 (1-911)

Qy 48 GluValLysAlaSerGlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGlu 67
Db 899 AAAATAATCGCGCGGAGACACATTTGGGACTAAATTTAGAGTAACTACATTGGTGAG 840
Qy 68 SerHisGlyGlyValGlyCysValIleSerGlyCysProArgIleProLeuThr 87
Db 839 TCACACGAGGTGGTGGTGTGTTGATAATTTGATGATGCTCTCGCGCATCCCCCTATCC 780
Qy 88 GluAlaAspLeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSer 107
Db 779 GAAGCTGATATGCAATTTGATCTAGATAGAGAGGCGCAGGTGAGCCGCAATTTACACT 720
Qy 108 ThrArgLysGluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThr 127
Db 719 CCAGAAAAGAGAGCGGATACCTGCAAAATATCTTCGGTGTGTTCTGAAGGACTGACTACT 660
Qy 128 GlyThrProIleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGlu 147
Db 659 GGGACGCCGATTCATGTTATTTGATCAAAATACTGATCAGAGAGGATTAGATTACAGTGAA 600
Qy 148 IleAlaAsnValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyVal 167
Db 599 ATGTCAGTTGCTTACAGCCCTTACATGACAGACCACTTATGACATGAAGTATGGTGT 540
Qy 168 ArgAlaValGlnGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAla 187
Db 539 AGATCACTTCAGGTGGGGGTAGATCTTCAGCAAGACAAACAATTTGAAGAGTGTGCT 480
Qy 188 GlyAlaLeuProLysLysIleLeuLysLeuLysCysGlyLeuGluLeuSerPheVal 207
Db 479 GGAGGTGTTGCTAAGAAAAATTTCTCAAACTATATGACAGAACTGAGATTCTTGCTTATGTC 420

QY 208 SerLysValHisGlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeu 227
 Db 419 TCTCAAGTCACCAAGTGTGTTACTTCCAGAGGTGTGTGATCACCACCTCTTACACTT 360
 QY 228 GluGlnIleGluSerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIle 247
 Db 359 GATCAGATGAGAGCAATATTGTTCAGATGTCAGATCCTGAATATGACAGAGATGATA 300
 QY 248 AspAlaIleAspArgValArgValArgGlyAspSerValGlyValIleThrCysVal 267
 Db 299 GCTGCCATTGATGCTGTCGAGTGAAGGGGATTCCTGTTGGTGT 255
 QY 268 AlaArgAsnValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeu 287
 Db 255 ----- 255
 QY 288 AlaLysAlaMetLeu-SerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAl 307
 Db 254 -----GTTGTCTACTACTGCAACAAAGGCGTTTGAATTTGGGAGTGGATTTC 207
 QY 307 aGlyThrAspLeuThrGlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySe 327
 Db 206 AGGCATCTCTTGACTGGGAGTGAACATAATGACGAGTCTTACACAGCAAAACATGGAAG 147
 QY 327 rValArgThrArgThrAsnArgSerGlyGlyValGlnGlyGlyIleSerAsnValGluI 347
 Db 146 AATCCGAACAAGACAAATCGTTCTGTGTAATACAGGAGGAGGAAATATCAATGGTGAAT 87
 QY 347 eValHisPheLysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrVa 367
 Db 86 TATAAATCATGAGTAATAGCTTTCAGGCAACATCTACATAGGAAAGAAAGCAACATACAGT 27
 QY 367 lSerArgGluArgGln 372
 Db 26 TACTAGAGAAAAA 11

RESULT 9
 LOCUS COL14396
 DEFINITION GR_EB015107.r GR_Eb Gossypium raimondii cDNA clone GR_EB015107
 3', mRNA sequence.
 ACCESSION COL14396
 VERSION COL14396.1 GI:48813083
 KEYWORDS EST.
 SOURCE Gossypium raimondii
 ORGANISM Gossypium raimondii

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
 1. (bases 1 to 865)
 Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
 Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
 Wing,R.A.
 Global assembly of Cotton ESTs
 Unpublished (2004)
 Contact: Rod A. Wing
 Arizona Genomics Institute
 The University of Arizona
 Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
 Tel: 520 626 9595
 Fax: 520 621 1259
 Email: http://genome.arizona.edu
 Placed: 015 row: 1 column: 07.
 Location/Qualifiers

FEATURES
 source
 1. .865
 /organism="Gossypium raimondii"
 /mol_type="mRNA"
 /db_xref="taxon:29730"
 /clone="GR_EB015107"
 /tissue_type="floral"
 /dev_stage=" 3 to +3 DPA"
 /lab_host="DH10B"
 /clone_lib="GR_Eb"

/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
 EcoRV; Library made by Invitrogen with RNA supplied by
 Wendle lab. Directional cloned into NotI-EV. Clones
 plated/picked by AGI. More glycerol clones held in -80."

ORIGIN
 Alignment Scores:
 Pred. No.: 6,53e-106 Length: 865
 Score: 1087.00 Matches: 200
 Percent Similarity: 88.0% Conservatives: 44
 Best Local Similarity: 72.20% Mismatches: 32
 Query Match: 49.41% Indels: 1
 DB: 7 Gaps: 0

US-10-677-179-8 (1-429) x COL14396 (1-865)

QY 48 GluValLysAlaSerGlyAsnThrPheGlyAsnTyr-PheGlnValAlaThrTyrGlyGl 67
 Db 37 CAGATACAGCAGCTGGGAGCAGATTTGGAATATATCTTCCGTGTACCAACATTTGGAGA 96
 QY 67 uSerHisGlyGlyValGlyCysValIleSerGlyCysProProArgIleProLeuTh 87
 Db 97 ATCTAATGGAGTGTGTGTGTGTATAGTGTGATGCTCTCTAGGATTCCTCTCTTC 156
 QY 87 rGluAlaAspLeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSe 107
 Db 157 AGAGCTGATTTGCAAGGTGATCTTGACCGAAGGCGGGTCAGAGCCGTGTACTAC 216
 QY 107 rThrArgLysGluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrTh 127
 Db 217 CCCAAGGAAAGAGACAGACATGCGGAATATATTTCTGGAGTTTCTGAAGGAGTGCATAC 276
 QY 127 rGlyThrProIleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGl 147
 Db 277 TGGAACCAATATACATGTACTTGTAACCAATCTGATCAGAGAACAGATTTACAAGGA 336
 QY 147 uIleAlaAsnValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyVa 167
 Db 337 AATGTCATAGCTTATAGCCATCTCATGCTGATGTCACCACTATGACATGAATATGGTGT 396
 QY 167 lArgAlaValGlnGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAla 187
 Db 397 CAGGCGAGTGCAGGCGTGTGCAGATCATCAGCAGAGAAACCATTTGAAGAGTTGCCTC 456
 QY 187 aGlyAlaLeuProLysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheVa 207
 Db 457 TGGAGCTATTGCTAAGAAAAATTTCTCAAGCAATTTTTCAGGAACTGAGGTTCTTGCATATGT 516
 QY 207 lSerLysValHisGlnValValLeuProGluAspAlaValAspTyrGlySerValThrLe 227
 Db 517 CTCTCAAGTTCAACAGGTTCTTCTTCCAGATGGGTGATGACCATGACATGACATGTAATCT 576
 QY 227 uGluGlnIleGluSerAsnIleValArgCysProAspProGluTyrAlaGluLysMetI 247
 Db 577 TGACCAGATAGAGCAATATTGTAAGGTGCCCAATCTCTGATACCTGAGAAATGAT 636
 QY 247 eAspAlaIleAspArgValArgValArgGlyAspSerValGlyGlyValIleThrCysVa 267
 Db 637 CGCTGCTATCGATGCTGTGCGGATAAGAGAGATTTCCGTTCGCGGTGTTCACCTGCATG 696
 QY 267 lAlaArgAsnValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLe 287
 Db 697 AGTGAGGAATGCTCCGCGTGGGCTTGGTTTCCAGGTTTTTGTATAGCTTTGAAGCAGAGCT 756
 QY 287 uAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAl 307
 Db 757 TGCTAAGGCTGTAATGTCTAGTACCTGCTGATCCAAAGGCTTTGAAGTGGAGTGGATTTTC 816
 QY 307 aGlyThrAspLeuThrGlySerGluHisAsnAspGluPheTyrMetAsp 323
 Db 817 AGGTGCTTTTCTAACCGGTAGTGAACATAATGACGAGTTCTTATACTGAT 865

RESULT 10


```

BG600177      BG600177      786 bp      mRNA      linear      EST 07-MAR-2003
LOCUS          EST505072 cSTS Solanum tuberosum cDNA clone cSTS28C6 5' sequence,
DEFINITION     mRNA sequence.
ACCESSION      BG600177
VERSION        BG600177.1 GI:13617313
KEYWORDS       EST.
SOURCE         Solanum tuberosum (potato)
ORGANISM       Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
REFERENCE      1 (bases 1 to 786)
AUTHORS        van der Hoeven R., Bezzerides J., Sun, H., Cho, J., Chiemiango, A.,
                Bougri, O., Buell, C.R., Roming, C., Tanksley, S. and Baker, B.
                Generations of ESTs from sprouting potato eyes
                Unpublished (2000)
                Contact: Robin Buell
                The Institute for Genomic Research
                9712 Medical Center Dr, Rockville, MD 20850, USA
                Email: potato-array@igr.org
                This clone can be obtained from the University of Arizona Genomics
                Institute. Orders can be made through URL:
                http://genome.arizona.edu/orders/
                Seq primer: M13P-R.
FEATURES       Location/Qualifiers
                source
                1..786
                /organism="Solanum tuberosum"
                /mol_type="mRNA"
                /cultiivar="Kennebec"
                /db_xref="taxon:4113"
                /clone="cSTS28C6"
                /tissue_type="sprouting eyes from tubers"
                /dev_stage="12-14 weeks post harvest"
                /lab_host="SOLR"
                /clone_lib="cSTS"
                /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
                XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
                taken from tubers. The tubers were incubated at 26C in the
                dark for 2-3 weeks prior to sprouting. The eyes were
                frozen in liquid nitrogen immediately upon removal from
                tubers."
ORIGIN
Alignment Scores:
Pred. No.:      3.9e-101      Length:      786
Score:          1042.00      Matches:     197
Percent Similarity: 89.49%      Conservative: 33
Best Local Similarity: 76.65%      Mismatches:  26
Query Match:    47.36%      Indels:      1
DB:              4              Gaps:        0
US-10-677-179-8 (1-429) x BG600177 (1-786)
QY      46 ArgLeuGluValLysAlaSerGlyValThrPheGlyAsnTyrPheGlnValAlaThrTyr 65
DB      13 CGACTAGAGATACAGCGCTGCTGGTAATACATTGGAAATATTTCGGCGGTACACATTTT 72
QY      66 GlyGluSerHisGlyGlyValGlyCysValIleSerGlyCysProProArgIlePro 85
DB      73 GGAGAAATCTCATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 132
QY      86 LeuThrGluAlaAspLeuGlnValGluLeuAspArgArgProGlyGlnSerArgIle 105
DB      133 CTCTCAGAGTCGAGTATCGAAGTGAACCTTGACAGGAGGAGCCAGGTCAAAGCCGAATT 192
QY      106 ThrSerThrArgLysGluThrAspThrCysValIleLeuSerGlyThrHisGlyVal 125
DB      193 ACCACCAAGAAAAGAGACTGATCTTGCAAAATATCATCAGGACCTGCCGAGTGTG 252
QY      126 ThrThrGlyThrProIleLeuValIleValProAsnThrAspGlnIleGlySerAspHis 145
DB      253 ACTACGGGATCTCCCAATCAAGTTGAAGTACCACACTGACCAGAGAGGAATGACTTAT 312

```

```

QY      146 ArgGluIleAlaAsnValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyr 165
DB      313 AGTGAATATCCCTTGGCTTATAGCCCACTCATGCTGATGCAACTTATGACTTCAAGTAT 372
QY      166 GlyValArgAlaValGlnGlyGlyValArgSerSerGlyArgLysThrValGlyArgVal 185
DB      373 GGAGTTAGATCAGTACAGGGGGGGGTAGATCTTCAGCAAGAGAAACCATCGGAGAGTC 432
QY      186 AlaAlaGlyAlaLeuProLysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSer 205
DB      433 GCTGCTGGAGCTGTTCAGAGAAAATCTCAAACTTTATTTCAGGAACCTGAGATCCTTGCT 492
QY      206 PheValSerLysValHisGlnValValLeuProGluAspAlaValAlaGlyThrGlySerVal 225
DB      493 TATGTTTTCACAAGTTCACAATGTTGACTTCCAGAGGATTTGGTTGATAACCCAGAGTGTG 552
QY      226 ThrLeuGluGlnIleGluSerAsnIleValArgCysProAspProGluTyrAlaGluLys 245
DB      553 ACCTAGAGCAGATTGAAGCAATATTTGCTTGCCTCCGACCTGATATATGCAGAAAG 612
QY      246 MetIleAspAlaIleAspArgValArgValArgGlyAspSerValGlyGlyValIleThr 265
DB      613 ATGATTGCTGCCATTGATTATGTACGAGTGAGGGGGGATCTGTTGGTGGTGTGCTAACT 672
QY      266 CysValAlaArgAsnValProArg-GlyLeuGlySerProValPheAspLysLeuGluSe 285
DB      673 TGCATTGTTAGAAATGTTCCACGAAGGTCTTGGTACACCATCTTTGATAAATTTGAAGC 732
QY      285 xGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPheGlu 301
DB      733 TGAGCTGGCTTAAGCTTGCATGTCATTCACGAGCACCACCAAGGGTTTTCAG 781
RESULT 11
BM411032      835 bp      mRNA      linear      EST 22-JAN-2002
LOCUS          EST585359 tomato breaker fruit Lycopersicon esculentum cDNA clone
DEFINITION     CLEGS5M18 5' end, mRNA sequence.
ACCESSION      BM411032
VERSION        BM411032.1 GI:18262662
KEYWORDS       EST.
SOURCE         Lycopersicon esculentum (tomato)
ORGANISM       Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE      1 (bases 1 to 835)
AUTHORS        Alcalá, J., Vrebalov, J., White, R., Vision, T., Karameycheva, S.A.,
                Tsai, J., Bougri, O., Kirkness, E., Utterback, T., Van Aken, S.,
                Roming, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and
                Giovannoni, J.
                Generation of ESTs from tomato fruit tissue, breaker stage (2002)
                Unpublished (2002)
                Contact: CUGI
                Clemson University Genomics Institute
                Clemson University
                100 Jordan Hall, Clemson, SC 29634, USA
                Email: http://www.genome.clemson.edu/orders/index.html
                This clone is available through the Clemson University Genomics
                Institute
                Seq primer: T3.
FEATURES       Location/Qualifiers
                source
                1..835
                /organism="Lycopersicon esculentum"
                /mol_type="mRNA"
                /cultiivar="TA496"
                /db_xref="taxon:4081"
                /clone="cLEGS5M18"
                /tissue_type="pericarp"
                /dev_stage="breaker"
                /lab_host="SOLR"
                /note="Vector: pBluescriptSKmCvadapt; Site_1: EcoRI;

```


[illegible]

RESULT 13					
CN144654					
LOCUS	CN144654	814 bp	mRNA	linear	EST 01-APR-2004
DEFINITION	WOUND1_23_E02.g1 A002 Wounded leaves Sorghum bicolor cDNA clone				
	WOUND1_23_E02 A002 5' , mRNA sequence.				

Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in the Human Genome Center, University of Tokyo Institute of Medical Science; plant material and RNA prepared at Texas A & M University; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.

ORIGIN

Alignment Scores:	
Pred. No.:	1,49e-99
Score:	1027.50
Percent Similarity:	86.96%
Best Local Similarity:	80.63%
Query Match:	46.70%
DB:	7
Length:	814
Matches:	204
Conservative:	16
Mismatches:	30
Indels:	3
Gaps:	2

US-10-677-179-8 (1-429) x CN144654 (1-814)

Qy	4	ValProIysProGlnGlnValAlaHisSerArgAlaArgLeuAlaProArgAlaIleGly	23
Db	62	GTGTCCGACGGCGCGGTGGCGCCAGGGCGTCCACCGGTTCTCCCGCGCGGATAGGC	121
Qy	24	AlaLeuLeuGluPheAlaProAlaSerSerSerLeuArgPheAlaVal--HisArgCys	42
Db	122	GCAGTCCGGAGTCCGCCCGCCGG-----TCGCTCGCTTCTTCGCTCGCGCCGCCCGCGC	175
Qy	43	ArgThrAlaArgLeuGluValLysAlaSerGlyAsnThrPheGlyAsnTyrPheGlnVal	62
Db	176	CGCGCGCGCTCGCTAGAGGTGAAGGCATCGGGAATGTGTTCGGGAATACTTTCAGGTT	235
Qy	63	AlaThrTyrGlyClnSerHisGlyGlyValGlyCysValIleSerGlyCysProPro	82
Db	236	GCAACCTATGGCGAATCCCATGGAGGGGTGTGGTGTGTATCATAGTGCCTGCCCAACC	295
Qy	83	ArgIleProLeuThrGluAlaAspLeuGlnValIglLeuAspArgArgCysProGlyGln	102
Db	296	AGAATTCTCTCTCAGGCGACATCGCAAGTAGAACTCGATAGAGACGTCGGGTCAA	355
Qy	103	SerArgIleThrSerThrArgLysGluThrAspThrCysLysIleLeuSerGlyThrHis	122
Db	356	AGTAGAATAACAACCCCAAGAAAGGAGACTGATACATGC AAAATTCATCAGGGACACAT	415
Qy	123	GluGlyValThrThrGlyThrProIleLeuValIleValProAsnThrAspGlnIleGly	142
Db	416	GATGGGATGACTACTGGCACACCAATTCACGCTCTTTGTCCCGAACACAGATCAAGAGGC	475
Qy	143	SerAspHisArgGluIleAlaAsnValTyrArgProSerHisAlaAspAlaThrTyrAsp	162
Db	476	GGTGATTACAGTGAATGGCTAAGGGGTACAGACCATCCCATCGAGATGCACCTATGAC	535
Qy	163	PheLysTyrGlyValArgAlaValGlnGlyGlyArgSerSerGlyArgLysThrVal	182
Db	536	TTCAAGTATGGTGTACAGAGCTGTGCAGGGAGGCGGAAGGTTCATCGGCCAGAGAAACCAT	595
Qy	183	GlyArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysLeuLysCysGlyLeuGlu	202
Db	596	GGCAGGGTGGCTCGAGAGCTCTTGCAAGAAAATCTGAAGCTCAAAATCAGGAGTGGAG	655
Qy	203	IleLeuSerPheValSerLysValHisGlnValValLeuProGluAspAlaValAspTyr	222
Db	656	ATCTTGGCATTGTCTTAAAGTGCATCAAGTCGTACTCCAGAAGATGCAGTTGATTAT	715
Qy	223	GlySerValThrLeuGluGlnIleGluSerAsnIleValArgCysProAspProGluTyr	242
Db	716	GAGACTGTAGCCTTGGNACAGATAGAGAGCAACATGTTAGATGTCTGTATCCAGATAT	775
Qy	243	AlaGluLysMetIleAspAlaIleAspArgValArgVal	255
Db	776	GCAGAGAAGATGATGTCTGCATTGTATAAAGTACGAGTT	814

RESULT 14
CF303160
LOCUS

723 bp mRNA linear EST 15-AUG-2003

DEFINITION ABFI-01-104.g1 ABF3-overexpressing transgenic rice lambda phage CDNA library (ABFI) Oryza sativa (japonica cultivar-group) cDNA clone ABFI-01-104, mRNA sequence.

ACCESSION CF303160.1 GI:33674921

VERSION

KEYWORDS

SOURCE

ORGANISM

Oryza sativa (japonica cultivar-group)

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaeae; Oryza.

REFERENCE 1 (bases 1 to 723)

AUTHORS Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C., Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.

TITLE Large-scale Sequencing Analysis of Rice ESTs

JOURNAL Unpublished (2003)

COMMENT Contact: Nahm B.H.
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division of Bioscience and Bioinformatics, Myongui University
Yongin, Gyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bhnahm@bio.com, bhnahm@bio.myongji.ac.kr.

FEATURES

Location/Qualifiers

1..723

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"

/cultivar="Nackdong"

/db_xref="taxon:39947"

/clone="ABFI-01-104"

/tissue_type="leaf"

/dev_stage="14 days after germination"

/lab_host="E.coli SOLR"

/clone_lib="ABF3-overexpressing transgenic rice lambda phage cDNA library (ABFI)"

/note="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2: XhoI; Leaf was dried for 2hrs. cDNA was inserted into lambda Uni-ZAP XR vector at 5' end with EcoRI and 3' end with XhoI site. mRNA was prepared from ABA-responsive element binding transcription factor 3 overexpression line."

ORIGIN

Alignment Scores:

Pred. No.: 4, 89e-99 Length: 723

Score: 1022.00 Matches: 199

Percent Similarity: 93.62% Conservative: 21

Best Local Similarity: 84.68% Mismatches: 14

Query Match: 46.45% Indels: 2

DB: 7 Gaps: 0

US-10-677-179-8 (1-429) x CF303160 (1-723)

QY 89 AlaAspLeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThr 108

DB 20 GCAGATATGCAAGTAGAAGTCTGACCGGAGACGCGCAGCAGCAGAGATTAACACCCCA 79

QY 109 ArgLysGluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGly 128

DB 80 AGAAAGGAGAGCTGACACTTGCAAAATCTTTCAGGGACACATGAAGGAATGACCACTGGG 139

QY 129 ThrProIleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIle 148

DB 140 ACACCAATTCATGTTTGTCCGACACACAGATCAGAGGGGGGTATCAGTGAATG 199

QY 149 AlaAsnValThrArgProSerHisAlaAspAlaThrThrAspPheLysThrGlyValArg 168

DB 200 GCTAAGGCTCAGACCTTCATGCAGATGCACTTATGACTTCAATACGGTGTAGA 259

QY 169 AlaValGlnGlyGlyLysSerSerGlyArgLysThrValGlyArgValAlaAlaGly 188

DB 260 GCAGTGCAGGGAGGTGGAAGATCATCAGCAAGAGAGACCATTTGGAAGGGTGGCTGCAGGA 319

QY 189 AlaLeuProLysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSer 208

DB 320 GCTCTTGCAAGAAAATTTCTTAAGCTCAATCTCGAGTAGAGATCTTGGCGTTGTGTCC 379

QY 209 LysValHisGlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGlu 228

DB 380 AAGGTGCATCAAGTTGTACTACCAAGATCCCGTTGATTATGACACTGTAAACAATGGAA 439

QY 229 GlnIleGluSerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAsp 248

DB 440 CAGATAGAAACAACATTTAGATGCTCTGATCAGAAATATGCACAGATGATGAT 499

QY 249 AlaIleAspArgValArgValArgGlyAspSerValGlyGlyValIleThrCysValAla 268

DB 500 GCAATCGATAAAGTACGAGTTAGAGTTGATTCGATTGGTGGTGCACATGCATTGC 559

QY 268 aArgAsnValProArgGlyLeuGlySerProValPheAspLysLeuGluLeuAla 288

DB 560 AAGAAATGTTCTCTCGTGGGATGGCGTCTCTGTTATTTGACAACTTGAGGCTGAATTGCC 619

QY 288 aLysAlaMetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaG1 308

DB 620 GAAAGCTATGCTTCTCTCTCGCAGCAGG-GGGTTTGAGATCGGCAGTGGATTGTCAGG 678

QY 308 YThrAspLeuThrGlySerGluHisAsnAspGluPheTyrMet 322

DB 679 TACTGACTACACTGGAAGTGAGCATATGATGAGTTCATATG 721

RESULT 15

CK279831

LOCUS

DEFINITION EST725909 potato abiotic stress cDNA library Solanum tuberosum cDNA clone POAEE54 5' end, mRNA sequence.

ACCESSION CK279831

VERSION CK279831.1 GI:39836809

KEYWORDS

SOURCE

ORGANISM

Solanum tuberosum (potato)

Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 755)

AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.

TITLE Generation of ESTs from abiotic stressed potato tissue

JOURNAL Unpublished (2003)

COMMENT Other_ESTs: EST725910
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics Institute via http://genome.arizona.edu/orders/ .
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

source

1..755

/organism="Solanum tuberosum"

/mol_type="mRNA"

/cultivar="Kennebec"

/db_xref="taxon:4113"

/clone="POAEE54"

/tissue_type="abiotic stress treated leaf and root tissue"

/lab_host="DH10B-Tona"

/note="vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering

applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:

Pred. No.: 5.2e-99 Length: 755
 Score: 1022.00 Matches: 188
 Percent Similarity: 91.39% Conservative: 35
 Best Local Similarity: 77.05% Mismatches: 21
 Query Match: 46.45% Indels: 0
 DB: 7 Gaps: 0

US-10-677-179-8 (1-429) x CK279831 (1-755)

Qy	46	ArgLeuGluValLysAlaSerGlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyr	65
Db	22	CCCTAGAGATACAGGCTGGTGATGATATTTGGAACTACTTCCGGTGAACAACTTTT	81
Qy	66	GlyGluSerHisGlyGlyValGlyCysValIleSerGlyCysProArgIlePro	85
Db	82	GGAGAACTCTCATGGTGGTGGAGTTGGTTGATTATTTGATGGATGTCCCCCGTCTCCCA	141
Qy	86	LeuThrGluAlaAspLeuGlnValGluLeuAspArgArgProGlyGlnSerArgIle	105
Db	142	CTTCTGAATCTGATATGATGAGTGAACCTTGACAGAGAGGCCAGGTCAAAGCGGATT	201
Qy	106	ThrSerThrArgLysGluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyVal	125
Db	202	ACCACACTAGGAAGAGACTGACACTTGCAAAATTTTCATCAGGCACCTGCAGATGGGCTG	261
Qy	126	ThrThrGlyThrProIleLeuValIleValProAsnThrAspGlnIleGlySerAspHis	145
Db	262	ACTACTGGATCTCCAACTCAAGTTGAAGTACCTTAACATCTGATCAGAGAGAAATGACTAT	321
Qy	146	ArgGluIleAlaAsnValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyr	165
Db	322	AGTGAATGTGCGTTGTACAGGCCATCTCATGCAGATGCCACTTATGACTTCAAGTAT	381
Qy	166	GlyValArgAlaValGlnGlyGlyArgSerSerGlyArgLysThrValGlyVal	185
Db	382	GGAGTGAGATCTGTACAGGGGGTGTGATGATCATCGACAGAGACCACTTGGGAGATT	441
Qy	186	AlaAlaGlyAlaLeuProLysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSer	205
Db	442	GCTGCTGGAGCAGTTGCTTAAGAAATTTCTCAAACTTATTCTGGAGCTGAGGTTCTTGCT	501
Qy	206	PheValSerLysValHisGlnValValLeuProGluAlaAspTyrGlySerVal	225
Db	502	TATGTTTCTCAAGTTCCCAAGTTGTACTTCTCTGAGGATTTGATTGATCATCAGAAATGTG	561
Qy	226	ThrLeuGluGlnIleGluSerAsnIleValArgCysProAspProGluTyrAlaGluLys	245
Db	562	ACTTAGAGCAGATAGAAAGCAATATTGTCGATGCCAGATCCGGAATATGCAGAGAG	621
Qy	246	MetIleAspAlaIleAspArgValArgGlyAspSerValGlyValIleThr	265
Db	622	ATGATTGCTGCTATTGATGCTGTACGAGTGAGAGGGGATTCTGTTGGTGGTGTGTAAC	681
Qy	266	CysValAlaAlaArgAsnValProArgGlyLeuGlySerProValPheAspLysLeuGluSer	285
Db	682	TGCATTGTTAGAAATCTCCACGGGGTCTTGTGTACACCACTGCTTCGATAAACTTGAAGCT	741

Qy 286 GluLeuAlaLys 289
 Db 742 GAGTAGCCANA 753

Search completed: August 27, 2005, 07:09:56
 Job time : 3372 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 27, 2005, 04:27:24 ; Search time 223 Seconds
(without alignments)

3147.815 Million cell updates/sec

Title: US-10-677-179-8

Perfect score: 2200

Sequence: 1 MTTVPKQVQVAHSRRARLAPR.....VAQCEMFALNTALQBPVGSF 429

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10677179/runat_25082005_150240_24862/app_query.fasta_1.583
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=p20 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=p20 -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10677179 @CNC_1_1_105 @runat_25082005_150240_24862 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

1: /cgn2_6/prodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2200	100.0	1626	4	US-09-743-207-7
2	1797.5	81.7	1635	4	Sequence 7, Appli
3	1257	57.1	966	4	Sequence 11, Appli
4	1066	48.5	1015	4	Sequence 5, Appli
5	859	39.0	1134	4	Sequence 5375, Ap
6	850	38.6	1089	4	Sequence 120, App
7	847	38.5	1239	3	Sequence 20, Appl
8	847	38.5	4530	3	Sequence 26, Appl
9	845	38.4	640681	4	Sequence 1, Appli
10	840.5	38.2	1830121	4	Sequence 1, Appli
11	840.5	38.2	1830121	4	Sequence 1, Appli
12	826.5	37.6	1080	4	Sequence 4099, Ap

13	800.5	36.4	1664976	4	US-08-916-421B-1	Sequence 1, Appli
14	800.5	36.4	1664976	4	US-09-692-570-1	Sequence 1, Appli
15	786	35.7	1140	4	US-09-252-991A-13588	Sequence 13588, A
16	786	35.7	1233	4	US-09-252-991A-13477	Sequence 13477, A
17	786	35.7	1263	4	US-09-252-991A-13857	Sequence 13857, A
18	756.5	34.4	6464	3	US-09-221-017B-168	Sequence 168, App
19	741	33.7	1104	4	US-09-328-352-3536	Sequence 3536, Ap
20	731	33.2	2312	4	US-09-103-331-1	Sequence 1, Appli
21	729	33.1	1125	4	US-09-540-236-310	Sequence 310, App
22	729	33.1	2410	4	US-09-596-002-12	Sequence 12, Appl
23	727	33.0	2312	4	US-09-631-594-45	Sequence 45, Appl
24	722.5	32.8	1053	3	US-09-610-040-1	Sequence 1, Appli
25	722.5	32.8	1053	3	US-09-610-040-7	Sequence 7, Appli
26	722.5	32.8	1053	4	US-10-267-763-1	Sequence 1, Appli
27	722.5	32.8	1053	4	US-10-267-763-7	Sequence 7, Appli
28	644	25.3	1837	4	US-09-103-331-3	Sequence 3, Appli
29	644	25.3	1837	4	US-09-631-594-54	Sequence 54, Appl
30	642.5	25.2	1230025	4	US-09-198-452A-1	Sequence 1, Appli
31	642.5	25.2	1230230	4	US-09-438-185A-1	Sequence 1, Appli
32	614.5	27.9	17245	4	US-09-902-540-1073	Sequence 1073, Ap
33	611	27.8	1020	4	US-09-902-540-9186	Sequence 9186, Ap
34	467	21.2	4403765	3	US-09-103-840A-2	Sequence 2, Appli
35	467	21.2	441529	3	US-09-103-840A-1	Sequence 1, Appli
36	459.5	20.9	1179	4	US-09-107-532A-1186	Sequence 306, App
37	437	19.9	1278	3	US-09-134-001C-306	Sequence 3562, Ap
38	437	19.9	2168	4	US-09-710-279-3562	Sequence 530, App
39	426	19.4	5532	4	US-08-956-171E-530	Sequence 13, Appl
40	426	19.4	5532	4	US-08-781-986A-530	Sequence 2027, Ap
41	424	19.3	541	4	US-09-743-207-13	Sequence 86, Appl
42	419.5	19.1	1176	4	US-09-107-433-2027	Sequence 882, App
43	419.5	19.1	19390	3	US-08-961-527-86	Sequence 1577, Ap
44	416.5	18.9	1167	4	US-09-583-110-882	
45	413.5	18.8	1188	4	US-09-134-000C-1577	

ALIGNMENTS

RESULT 1
US-09-743-207-7
; Sequence 7, Application US/09743207
; Patent No. 6653531
; GENERAL INFORMATION:
; APPLICANT: Calico, Rebecca E.
; APPLICANT: Calico, Rebecca E.
; APPLICANT: Pember, Stephen O.
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-A
; CURRENT APPLICATION NUMBER: US/09/743,207
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: PCT/US99/16353
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 1626
; TYPE: DNA
; ORGANISM: Zea mays
US-09-743-207-7

Alignment Scores:
Pred. No.: 1.72e-240 Length: 1626
Score: 2200.00 Matches: 429
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-677-179-8 (1-429) x US-09-743-207-7 (1-1626)

Qy 1 MetThrValProLysProGlnValAlaHisSerArgAlaArgLeuAlaProArg 20


```
Db 418 GATGGGATGACTACTGGTACACCAATTCACGTCTCTTTGTCCTCCAAACACAGATCAAAAGGGT 477
Qy 143 SerAspHisArgGluLeuAlaAenValTyrArgProSerHisAlaAspAlaThrTyrAsp 162
Db 478 GGTGATTACGTGAAATGCTTAAGGGGTACAGACCATCCATCGATGCAACCTATGAC 537
Qy 163 PheLysTyrGlyValArgAlaValGlnGlyGlyArgSerSerGlyArgLysThrVal 182
Db 538 TTCAGTATGAGCTTAGAGCTGTGCAGGAGGTGGAAGGTTCATCAGCCAGAGAAACCAAT 597
Qy 183 GlyArgValAlaAlaGlyAlaLeuProLysLysLysLysLysLysLysLysLysLys 202
Db 598 GGCAGGGTGGCTGCAGAGAGCTCTTGAAGAAATTCCTAAAGCTCAAAATCAGAGTGGAG 657
Qy 203 IleLeuSerPheValSerLysValHisGlnValValLeuProLysAlaValAspTyr 222
Db 658 ATCTTGCCATTTGTTCTTAAGTGCACCAAGTCTTCTCCAGAGATGAGTGAATAT 717
Qy 223 GlySerValThrLeuGluGlnIleGluSerAsnIleValArgCysProAspProGluTyr 242
Db 718 GAGACTGTAACTTGGAAACATATAGAGAGCAACATCGTTAGATGCTCTGATCCAGATAT 777
Qy 243 AlaGluLysMetIleAspAlaIleAspArgValArgValArgGlyAspSerValGlyGly 262
Db 778 GCAGAGAAGATGATTCTGCTCCATTTGATACGGTACGAGTTAGAGGAGATTCATTTGGTGG 837
Qy 263 ValIleThrCysValAlaArgAsnValProArgGlyLeuGlySerProValPheAspLys 282
Db 838 GTGCTCATCATGATTCGCAAGAAATGTTCTCGTGGTCTTGGCTCTCTGTTTTCAGAAA 897
Qy 283 LeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPheGluIle 302
Db 898 CTTGAAGCTGAATGCAAAAGCCATGCTTCTCTCTGCAAGCAAGGGGTTTGAGAT 957
Qy 303 GlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPheTyrMet 322
Db 958 GGCAGTGGGTTCGCTGCTACGCACTTTACTGGAAGTGAAGCATAATGATGATGTTCTATATG 1017
Qy 323 AspLysAlaGlySerValArgThrArgThrAsnArgSerGlyGlyValGlnGlyGlyIle 342
Db 1018 GATGAGCTGGAAATGTGAGGACACCACTTAATCGCTCAGGCGGTGTTTCAAGGAGGATA 1077
Qy 343 SerAsnValGluIleValHisPheLysValAlaPheLysProThrProSerIleGlyVal 362
Db 1078 TCAATGGTGAATATTATTACTTCAAGTGGCTTTTAAAGCAACAGCACTATCGGAAG 1137
Qy 363 LysGlnAsnThrValSerArgGluArgGlnAsnValGluLeuAlaArgGlyArgHis 382
Db 1138 AAGCAAAATACTGTGTCAAGGAGCATGAGGATGTTGAACCTTTTGGCAAGGGGGGCCAT 1197
Qy 383 AspProCysValAlaProArgAlaValProValValGluSerMetAlaAlaLeuValLeu 402
Db 1198 GACCCCTGTGTGCTCCTCGAGCTGTCTATGTTGGATTCATGCTGCTGCTGCTGCTG 1257
Qy 403 MetAspGlnLeuMetAlaHisValAlaGlnCysGluMetPheAlaLeuAsnThrAlaLeu 422
Db 1258 ATGGACCACTCATGGCGATATGCCCAGTGTGAGATGTTTCCGCTGACCTTGGCCTA 1317
Qy 423 GlnGluProValGlySer 428
Db 1318 CAAGAGCCCATTTGGCTCT 1335
```

RESULT 3

```
US-09-743-207-11
; Sequence 11, Application US/09743207
; Patent No. 6653531
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Pember, Stephen O.
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-A
; CURRENT APPLICATION NUMBER: US/09/743,207
```

```
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: PCT/US99/16353
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 966
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-743-207-11

Alignment Scores:
Pred. No.: 2,318-133 Length: 966
Score: 1257.00 Matches: 245
Percent Similarity: 89.30% Conservative: 22
Best Local Similarity: 81.94% Mismatches: 26
Query Match: 57.14% Indels: 6
DB: 4 Gaps: 2
```

US-10-677-179-8 (1-429) x US-09-743-207-11 (1-966)

```
Qy 15 AlaArgLeuAlaProArgAlaIleGly-----AlaLeuLeuGluPheAlaPro 30
Db 75 GCGCGGCTCTCCCGCGCGCGCGCGCGGTTCCGCGCTTCCCGAGTCCGCGCG 134
Qy 31 AlaSerSerSerLeuArgPheAlaValHisArgCysArgThrAlaArgLeuValLys 50
Db 135 GCT-----TCCTCCGCTTCTCGTCCGCGCGCGCGCGCTCGCTAGAGTGAAG 188
Qy 51 AlaSerGlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGly 70
Db 189 GCGTCTGCAATGTAATTTGGGAACCTACTTCAGGTTTGAACCTTATGGAGAGTCTCATGA 248
Qy 71 GlyGlyValGlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAsp 90
Db 249 GCGGTGTTGGTGGTAATCAGTGGATGCCACCCAGCAATCCCACTTACTGAAGCAGAT 308
Qy 91 LeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLys 110
Db 309 ATGCAAGTAGAATCTGACCGAGACGCGCAGCAGCAGAGCAGATAACCCCAAGAAAG 368
Qy 111 GluThrAspThrCysLysIleLeuSerGlyThrHisGlyGlyValThrGlyThrPro 130
Db 369 GAGACTGACCTTGCAAAATTTCTTTCAGGACACATGAAGGAATGACCCTGGGACACA 428
Qy 131 IleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsn 150
Db 429 ATTCAATGTTTGTCCCGAACACACAGATCAGAGAGGGGGTGTATACAGTGAATGCTAAG 488
Qy 151 ValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaVal 170
Db 489 GCCTACAGACCTTCACATGCGAGTGCACCTTATGACTTCAAAATACGTTGTAGAGCAGTG 548
Qy 171 GlnGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeu 190
Db 549 CAGGAGGTGGAAGATCATCAGAGAGACCATTTGAAGGGTGGCTGCGAGGAGCTCTT 608
Qy 191 ProLysLysIleLeuLysLysLysCysGlyLeuGluIleLeuSerPheValSerLysVal 210
Db 609 GCMAAGAAATTCCTAAGCTCAATCTGGAGTAGAGATCTTGGCGTTTGTGTCCAAGGTG 668
Qy 211 HisGlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIle 230
Db 669 CATCAAGTTGTACTACCAAGAGATGCCGTTGATTATGACATCTGTAACCAATGGAAACAGATA 728
Qy 231 GluSerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIle 250
Db 729 GAAAGCAACATGTTAGATGCTCTGATCCAGAAATATGCAGAGAGATGATTGATGACTC 788
Qy 251 AspArgValArgValArgGlyAspSerValGlyGlyValIleThrCysValAlaArgAsn 270
```

Db 789 GATAAAGTACGAGTTAGAGTGATTCGATTGGTGGTGCATGCATGCAATGCAAGAAAT 848
Qy ValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAla 290
Db 849 GTTCTCGGGATTGGCTCTCTCTATTATGCAAACTTGAGGCTGAAATGGCGAAGCT 908
Qy 291 MetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThr 309
Db 909 ATGCTTTCTCTCTCTGCAAGCAAGGGGTTTGAGATCGGCGATGGGATTTGTGTTCACT 965

RESULT 4

US-09-743-207-5
; Sequence 5, Application US/09743207
; Patent No. 6653531
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Pember, Stephen O.
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-A
; CURRENT APPLICATION NUMBER: US/09/743,207
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: PCT/US99/16353
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1015
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-743-207-5

Alignment Scores:
Pred. No.: 1,48e-111 Length: 1015
Score: 1066.00 Matches: 201
Percent Similarity: 93.83% Conservative: 27
Best Local Similarity: 82.72% Mismatches: 15
Query Match: 48.45% Indels: 0
DB: 4 Gaps: 0

US-10-677-179-8 (1-429) x US-09-743-207-5 (1-1015)

Qy 186 AlaAlaGlyAlaLeuProLysLysIleLeuLysCysGlyLeuGluIleLeuSer 205
Db 8 GCTGCAGGAGCTGTTGCAAGAAAAATCTTAAGCTGAAATGGGAGTAGAGATTCTAGCA 67
Qy 206 PheValSerLysValHisGlnValValLeuProGluAspAlaValAspTyrGlySerVal 225
Db 68 TTTGTTTCCAAAGTCATCAAGTGGTACTTCTCTGAAGCGCAGTTGATATGAAACTCTT 127
Qy 226 ThrLeuGluGlnIleGluSerAsnIleValArgCysProAspProGluTyrAlaGluLys 245
Db 128 ACCCTGGATCAGATAGAGAGCAACATTTGTAGATGTCCTGATCCAGATATATGCAGAAG 187
Qy 246 MetIleAspAlaIleAspArgValArgValArgGlyAspSerValGlyValIleThr 265
Db 188 ATGATTGATGCAATTTGATTAAGTACGAGTTAATGGGAATTCGATTGGTGGGTGTCACA 247
Qy 266 CysValAlaArgAsnValProArgGlyLeuGlySerProValPheAspLysLeuGluSer 285
Db 248 TGCATTGCCAAGAAATGTTCTCTGTGGGCTTGAGCTCTCTGTTATTTGACAAACTGAGCT 307
Qy 286 GluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGly 305
Db 308 CTACTGGCAAGAGCTATGCTTCTCTCTGCAAGCAAGGGGTTTGAGATCGGTAGTGA 367
Qy 306 PheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPheTyrMetAspLysAla 325
Db 368 TTTGCAGGTAAGTACCTAAGTGAAGTGAACATGAGTGTCTTATATGACAGGCT 427
Qy 326 GlySerValArgThrArgThrAsnArgSerGlyGlyValGlnGlyIleSerAsnVal 345

Db 428 GGAATGTAAGAACACCAATCGCTCGGGCGGTCTACAGGGAGGATATCAATGGT 487
Qy 346 GluIleValHisPheLysValAlaPheLysProThrProSerIleGlyValLysGlnAsn 365
Db 488 GAAACTATATACTTCAAAAGTAGCTTTCAAGCCAACAGCAACTATTGGGAAGACAAAAT 547
Qy 366 ThrValSerArgGluArgGlnAsnValGluLeuLeuAlaArgGlyArgHisAspProCys 385
Db 548 ACTGTAAACAAAGGATCATGAGGATATCGAACTTCTGACAAAGGGGTGCGCATGACCCATGT 607
Qy 386 ValAlaProArgAlaValProValValGluSerMetAlaAlaLeuValLeuMetAspGln 405
Db 608 GTCTGCTCCTCGGGCTGTTCCAAATGCTGAGACGATGGCTGATTGGTCTCATGACACAG 667
Qy 406 LeuMetAlaHisValAlaGlnCysGluMetPheAlaLeuAsnThrAlaLeuGlnGluPro 425
Db 668 CTGATGGCACATGTTGCTCAGTGGGAGATGTTCCCGCTGAACCTCGCCCTACAAGAACCA 727
Qy 426 ValGlySer 428
Db 728 ATCGGCTCC 736

RESULT 5

US-09-489-039A-5375
; Sequence 5375, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709,2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5375
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5375

Alignment Scores:
Pred. No.: 6,94e-98 Length: 1134
Score: 859.00 Matches: 180
Percent Similarity: 64.99% Conservative: 65
Best Local Similarity: 47.75% Mismatches: 112
Query Match: 39.05% Indels: 20
DB: 4 Gaps: 6

US-10-677-179-8 (1-429) x US-09-489-039A-5375 (1-1134)

Qy 40 HisArgCysArgThrAlaArgLeuGluValLysAlaSerGlyAsnThrPheGlyAsnTyr 59
Db 16 CACGCCAACATAACGAAAAAGATTGGAGCCGTGATGGCAGGAATACATAATTGGACAACTC 75
Qy 60 PheGlnValAlaThrTyrGlyGluSerHisGlyGlyValGlyCysValIleSerGly 79
Db 76 TTTCCGCTCACTACCTCTCGGCGAATCGCACGCTCTGGCGCTGCGGTGATCGTTGACGGC 135
Qy 80 CysProProArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArgArgArg 99
Db 136 GTGGCCCGCGGCATCCCGCTACCCAGGCCACCTGACGACGACCTGATCGATCGTCTGCGC 195
Qy 100 ProGlyGlnSerArgIleThrSerThrArgLysGluThrAspThrCysLysIleLeuSer 119
Db 196 CCGGGTACTCTCGCTTACACCACCGCAGCGTTCGCGAGCGGATCAGGTCAAAATCTCTCTCC 255
Qy 120 GlyThrHisGluGlyValThrThrGlyThrProIleLeuValIleValProAsnThrAsp 139
Db 256 GCGCTCTTCGAAGCGGTTACCCCGGACCAAGTATCGGTCTGCTGATTGAGAACACCGAT 315


```
QY 140 GlnIleGlySerAspHisArgGluIleAlaAsnValTyrArgProSerHisAlaAspAla 159
Db 316 CAGCGTTCCGAGGATTACCGCGCGATCAAGAGATGTTCCGTCCGGGCCATCGCGATTAC 375
QY 160 ThrTyrAspPheLysTyrGlyValArgAlaValGlnGlyGlyArgSerSerGlyArg 179
Db 376 ACTACGAGCAAAAGATGTTCTCGCGCATATATCGCGCGCGGTGCTTCTTCTGCGCGC 435
QY 180 LysThrValGlyArgValAlaAlaGlyAlaLeuProLysLysLysLysLysCys 199
Db 436 GAAACCGCGATGCGGTAGCGCGCGCGCGGCGATTCGCAAAATACCTTCGCGCGAAGTTC 495
QY 200 GlyLeuGluIleLeuSerPheValSerLysValHisGlnValValLeuProGluAlaAspAla 219
Db 496 GGCATCGTCAATTCGCGCGTGCCTGACGCGCATGGCGGATATT-----CGCTGGCCATC 549
QY 220 ValAspTyrClySerValThrLeuGluGlnIleGluSerAsnIleValArgCysProAsp 239
Db 550 AAGACTGG-----GATCAGGTTGACGAGAACCCCTTCTTCTGTCGCGAC 594
QY 240 ProGluTyrAlaGluLysMetIleAspAlaIleAspArg-----ValArgVal 255
Db 595 CCGGAT-----AAATTCGACGCCCTCGATGAACTGATGCGCGCCCTGAAAAAA 642
QY 256 ArgGlyAspSerValGlyGlyValIleThrCysValAlaArgAsnValProArgGlyLeu 275
Db 643 GAGGGGACTCCATTCGCGCGAAGTGACGGTGTGCTGACGCGGTACCGCGACGCTC 702
QY 276 GlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMetLeuSerIlePro 295
Db 703 GGTGAGCCGCTTTCACCGCTGATGCGCGATATCGCCACGCGCTGATGAGCATCAAC 762
QY 296 AlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeuThrGlySerGlu 315
Db 763 GCGTGAAAGGGTGAATATGCGACGCGCTTCGAGGTGTAAGTTCGCGCGACGCGAA 822
QY 316 HisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThrArgThrAsnArgSer 335
Db 823 AACCGCGATGAG-----ATCACTAAGCGCGG-----TTCCAGACCAACCGCGC 867
QY 336 GlyGlyValGlnGlyIleSerAsnValGluIleValHisPheLysValAlaPheLys 355
Db 868 GCGCGCATCTTTCGCGCGCATCAGCAGCGCGCGCAGCATGCTTGCACCAACATCGCGCTGAAG 927
QY 356 ProThrProSerIleGlyValLysGlnAsnThrValSerArgGluArgGlnAsnValGlu 375
Db 928 CTTACTCCGATATTACCGTTCCTCCGTCATACGATTAATTCGCTTTTGGCGAAGTGGAG 987
QY 376 LeuLeuAlaArgGlyArgHisAspProCysValAlaProArgAlaValProValValGlu 395
Db 988 ATGATCAACCAAGGGCGTCACGATCCGTGCGTCGCGCATCGCGCGCTACCGATCGCTGAG 1047
QY 396 SerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisValAlaGln 412
Db 1048 GCGATGCTCGCGCATCTTCTGATGATCATCTTATTCGTCGTCAGCGCGCGCGAG 1098

RESULT 6
US-09-543-681A-120
; Sequence 120, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 120
; LENGTH: 1089
; TYPE: DNA
```

```
i ORGANISM: Proteus mirabilis
US-09-543-681A-120

Alignment Scores:
Pred. No.: 6,91e-87 Length: 1089
Score: 850.00 Matches: 167
Percent Similarity: 66.02% Conservatives: 72
Best Local Similarity: 46.13% Mismatches: 111
Query Match: 38.64% Indels: 12
DB: 4 Gaps: 3

US-10-677-179-8 (1-429) x US-09-543-681A-120 (1-1089)

QY 52 SerGlyAsnThrPheClyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGlyGly 71
Db 7 GCGGGAACAGTATCGGCAATATTATTAGAGTAACCACTTTTGGTGAGTCTCATGGCACA 66
QY 72 GlyValGlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAspLeu 91
Db 67 GCGTTAGGTTGCATTTGGTGTCTCTCCCGGACTACCTTTAACGCAAGCGGATCTA 126
QY 92 GlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLysGlu 111
Db 127 CAAAGTTGATTAGATAGACGTAACCCGGAACTTCCAGCTTATACCAACACAGCTAGAG 186
QY 112 ThrAspThrCysLeuIleLeuSerGlyThrHisGlyValThrThrGlyThrProIle 131
Db 187 CCGTATCAAGTCGGTATTTATCGGGTGTCTTTAATGGTGAACACCGGAACCAAGTATT 246
QY 132 LeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnVal 151
Db 247 GGAATTATTAAATAGAAATACCGATCAGCGCTCTCAAGATTATAGCGAAATTAAGATGTA 306
QY 152 TyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGln 171
Db 307 TTCCGCCCGGCGCATCGACTACACCTATGAACAGAAATATGGTTTACGTGATTATCGT 366
QY 172 GlyGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyValAlaLeuPro 191
Db 367 GCGCGGGGACGCTCTTCTGTCGAGAAACCGCATGCGTGTGCGACGGGTGCTATTGCC 426
QY 192 LysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysValHis 211
Db 427 AAAAATAACCTTAAACAAAAATTTGTTATGAAGTAAAGGTACTTATCTCAATTAGGG 486
QY 212 GlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGlu 231
Db 487 CCGATTAGTTGT-----GAGTTAGTTGATTGTTCAATTGTA-----GAA 525
QY 232 SerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIleAsp 251
Db 526 ACCAACCCATCTTTTTCGCCAGATCCCTTCGCGCTAGATGCCCTTGATGAATATATGCGA 585
QY 252 ArgValArgValArgGlyAspSerValGlyGlyValIleThrCysValAlaAlaArgAsnVal 271
Db 586 GCGCTAAAAAAGAGGGTAATCTATAGTGCCAAAGTTACCGTGTTCGCAAGGTGTA 645
QY 272 ProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMet 291
Db 646 CCGCAGGATTAGGTGAACCCGCTTTTGTATAGACTTGTATGCGGATTTGGCTCATGCGTTA 705
QY 292 LeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeu 311
Db 706 ATGAGCATCAACCGCATGAAAGCGCTTGAATTCGAGATGGTTTTTGATGTGGTGACCTTGA 765
QY 312 ThrGlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThrArg 331
Db 766 AAAAGGACTGAAACCGCGATGAAATTAACCAAGAGGATTTAGCAGT-----813
QY 332 ThrAsnArgSerGlyGlyValGlnGlyIleSerAsnValGluIleValHisPheLys 351
Db 814 ---AATCAGCGGGGGCGGTATTAGTGGCGCATTTAGTAGTGGTCAACCTTATTATTACACAC 870
```


Db 1159 ACGCAACGGCGCAA 1174

RESULT 8

US-09-064-693A-26

Sequence 26, Application US/09064693A

Patent No. 6210937

GENERAL INFORMATION:

APPLICANT: Ward, Thomas E.

TITLE OF INVENTION: DEVELOPMENT OF GENETICALLY

TITLE OF INVENTION: ENGINEERED BACTERIA FOR PRODUCTION

TITLE OF INVENTION: OF A SPECIFIC PLASTICS PRECURSOR

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: W. Gary Goodson

ADDRESSEE: INEL-Lotheed Martin Idaho

ADDRESSEE: Technologies Co.

STREET: P.O. Box 1625

CITY: Idaho Falls

STATE: Idaho

COUNTRY: USA

ZIP: 83415-3810

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage

COMPUTER: Toshiba Satellite Pro T2150CDS

OPERATING SYSTEM: Windows95

SOFTWARE: Word Perfect 7.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/064,693A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: W. Gary Goodson

REGISTRATION NUMBER: 22,387

REFERENCE/DOCKET NUMBER: LIT-PI-296

TELEPHONE: (208)526-9469

TELEFAX: (208)526-8339

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 4530 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-09-064-693A-26

Alignment Scores:

Pred. No.:	1.46e-85	Length:	4530
Score:	847.00	Matches:	178
Percent Similarity:	65.30%	Conservative:	61
Best Local Similarity:	48.63%	Mismatches:	106
Query Match:	38.50%	Indels:	21
DB:	3	Gaps:	5

US-10-677-179-8 (1-429) x US-09-064-693A-26 (1-4530)

QY 52 SerGlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGlyGly 71

Db 127 GCTGGAACACCAATTCGACCACTCTTTCGCGTAACCACTTCGGCGAATCGCAGCGGTG 186

QY 72 GlyValGlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaSerLeu 91

Db 187 GCGCTCGGTCGATCGTCGATGTTCCGCCAGGCAATTCGCTGACGGAAGCGACCTG 246

QY 92 GlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLysGlu 111

Db 247 CAACATGACCTGACCGCTGTCGCTCGGACATCGCGCTATACCAACCGAGCGCGAG 306

QY 112 ThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThrPro 131

Db 307 CCGGATCAGGTCAAATCTCTCCGGTGTTTGAAGCGGTACTACCGGCAACGACATT 366

QY 132 LeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaVal 151

Db 367 GCGTGTGTGATCGAAACACTGACGCGCTCTCAGGATTACAGTGGCGATTAGGACGT 426

QY 152 TyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGln 171

Db 427 TTCCGTCAGGCGCATGCGGATTACACCTACGAAACAAATAATCGTCTGCGGATTATCGC 486

QY 172 GlyGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuPro 191

Db 487 GCGCGTGACGCTTCTTCCGCCGGAACCGCATCGCGTGGCGGACGAGCTATTGCC 546

QY 192 LysLysIleLeuLysLysCysGlyLeuGluIleLeuSerPheValSerLysValHis 211

Db 547 AAAAATATCTCGCGAGAAATTTGGTATTTGAAATCCGTGGCTGCTGACCCAGATGGCG 606

QY 212 GlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGlu 231

Db 607 GACATT-----CGCTGATATCAAGACTGG-----TCGCAAGTTCGAG 645

QY 232 SerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaLeu 251

Db 646 CAAAATCCGCTTTTTCGCCGACCCGAC-----AAAATCGACCGCTTAGAC 693

QY 252 Arg-----ValArgValArgGlyAspSerValGlyValIleThrCysVal 267

Db 694 GAGTTGATCGCGCTGAAAGAGGCGGACTCCATCGGCGTAAAGTCAACCGTTGT 753

QY 268 AlaArgAsnValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeu 287

Db 754 GCCAGTGGCTTCTCGCGGACTTCGCGAGCGGCTTTTGACCGCTGGATCGACATC 813

QY 288 AlaLysAlaMetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAla 307

Db 814 GCCCATGCGCTGATGAGCATCAACGCGGTGAAAGCGTGGAAATTTGGCGACGCTTTCAC 873

QY 308 GlyThrAspLeuThrGlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySer 327

Db 874 GTGTGGCGCTGCGCGGACGCGAAGCCGCGATGAATCACCAGAAC-----921

QY 328 ValArgThrArgThrAsnArgSerGlyGlyValGlnGlyGlyIleSerAsnValGluIle 347

Db 922 ---GGTTTCCAGAGCAACCATCGCGGCGGCATTCTCGCGGTATCAGCAGCGGCGCA 978

QY 348 ValHisPheLysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrVal 367

Db 979 ATCATTGCCCATATGCGCGCTGAAACCGACCTCCAGCATTTACCGTCCGCGGTGATC 1038

QY 368 SerArgGluArgGlnAsnValGluLeuAlaArgGlyArgHisAspProCysValAla 387

Db 1039 AACCGCTTTGCGGAAGAAGTGGATGATGATACCAAGGCGGTCAGATCCCTGTGTGGG 1098

QY 388 ProArgAlaValProValValGluSer-MetAlaAlaLeuValLeuMetAspGlnLeuMe 407

Db 1099 ATCCGCGCAGTGGCGATCGCAGAGCGAATGCTGCGGATCGTTTAAATGGATCACC 1158

QY 407 talahisValAlaGln 412

Db 1159 ACGCAACGGCGCAA 1174

RESULT 9

US-09-790-988-1/c

Sequence 1, Application US/09790988

Patent No. 6632935

GENERAL INFORMATION:

APPLICANT: SHIGENOBU, SHUJI

APPLICANT: WATANABE, HIDEMI

APPLICANT: HATTORI, MASAHIRA

APPLICANT: SAKAKI, YOSHIYUKI

TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS

FILE REFERENCE: 081356/0159

CURRENT APPLICATION NUMBER: US/09/790,988

```
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Alignment Scores:
Pred. No.: 6,5e-82 Length: 640681
Score: 845.00 Matches: 172
Percent Similarity: 64.84% Conservatives: 64
Best Local Similarity: 47.25% Mismatches: 116
Query Match: 38.41% Indels: 12
DB: 4 Gaps: 3

US-10-677-179-8 (1-429) x US-09-790-988-1 (1-640681)

Qy 50 LysAlaSerGlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHis 69
Db 102991 AAAATGCTCTGGAATACAAATTCGGGAAAATATTCGTGTAAACCACTTTTGGCGAGTCACAC 102932

Qy 70 GlyGlyGlyValGlyCysValIleSerGlyCysProProArgIleProLeuThrGluAla 89
Db 102931 GAGAAAGCAATAGGGTGCATAATGACGGAACACCTCTCGGTCTTGAATATTCCTGTAAA 102872

Qy 90 AspLeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArg 109
Db 102871 GATTTGCAATATGATTTGAATCGTAGAGACCAAGGACTTCCCGTTACACGACTTTGGCG 102812

Qy 110 LysGluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThr 129
Db 102811 CGTGAACCTGTAGAGTAAATATATCTTCCGGTATATTTAAACGGGGTCACAACTGGTACT 102752

Qy 130 ProIleLeuValIleValProLeuThrAspGlnIleGlySerAspHisArgGluIleAla 149
Db 102751 AGTATGGTTTAAATATTTATATCATGATCATAGTCTCAAGACTATAGCATATTAAG 102692

Qy 150 AsnValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAla 169
Db 102691 AATTTATTTTCGACCGGACATCCGATTTACTTATGAAAAAATATGGAATTAGAGAT 102632

Qy 170 ValGlnGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAla 189
Db 102631 TATCGGGGGAGGTAGATCTTCTGCTCGGAAACTGCTATGAGAGTTGACGACGAGCT 102572

Qy 190 LeuProLysLysIleLeuLysLysCysGlyLeuGluIleLeuSerPheValSerLys 209
Db 102571 ATAGCAAAAAATACCTTAATGAGAAATATGGGATAACTATTTCGACATATTGTGACGA 102512

Qy 210 ValHisGlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGln 229
Db 102511 ATGGCAATATCAAAATGCCCT-----TTTAAATCTTGGCAAGA 102473

Qy 230 IleGluSerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAla 249
Db 102472 GTCGAAAAATATCTTTTCTGTTCTGCTCTCGAAAAAATTTTAGCACTAGAAAAATTTA 102413

Qy 250 IleAspArgValArgValArgGlyAspSerValGlyValIleThrCysValAlaArg 269
Db 102412 ATTAATATCTCAAAAAATAGTGTATTCGATTTGGTGGTGAATATACAAATTAATGCTGNA 102353

Qy 270 AsnValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLys 289
Db 102352 AATATCCCTGTAGGACTTGGAGAACCCAGTTTTTGTATCGCTTGATGCTGATTTATCACAC 102293

Qy 290 AlaMetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThr 309
Db 102292 GCCTTGATGATTAATATCGCGCAAAAGAGTAGAAATTTGGAGACGGTTTTTTCAGTAATA 102233

Qy 310 AspLeuThrGlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArg 329
Db 102232 AATCAACGAGGAGTGAACATCGTATGAA-----ATTACGCCACAGGATTTTAA--- 102182

Qy 330 ThrArgThrAsnArgSerGlyValGlnGlyGlyIleSerAsnValGluIleValHis 349
Db 102181 -----ACTAATCATCTCTGGTGGTATTCTAGTGGTATTAGTATGTCGAGAAATTTGTA 102128

Qy 350 PheLysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArg 369
Db 102127 CTAAAGTAGCATTTAAACCTACATCAATGATTCGAAAGCAGGCAATACATAAATAAA 102068

Qy 370 GluArgGlnAsnValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaProArg 389
Db 102067 AATRAACGAAAAAGTTCAAATAGTTACTTAAGGACGACATGATCCATGTAGTGTACGC 102008

Qy 390 AlaValProValValGluSerMetAlaLeuValLeuMetAspGlnLeuMetAlaHis 409
Db 102007 GCTGTTCCATCACTGAAGCAATGTTAGCAATTTGATTAATGATCATTTTATTAGATTTT 101948

Qy 410 ValAlaGlnCys 413
Db 101947 CGCGACAATGC 101936

RESULT 10
US-09-557-884-1
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,884
; FILING DATE: 25-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: JUN-5-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PBI86P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Alignment Scores:
Pred. No.: 1.12e-80 Length: 1830121
Score: 840.50 Matches: 166
Percent Similarity: 64.74% Conservatives: 69
```

Best Local Similarity: 45.73% Mismatches: 121
Query Match: 38.20% Indels: 7
DB: Gaps: 3

US-10-677-179-8 (1-429) x US-09-557-884-1 (1-1830121)

```
QY 52 SerGlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGlyGly 71
DB 211427 GCTGGTAATACAACTTTCCTGCTGAGCAACCTTGGAGAGTCACATGGTATT 211486

QY 72 GlyValGlyCysValIleSerGlyCysProAspGlyLeuThrGluAlaLeu 91
DB 211487 GCATAGGCTGATCTGATGCGCGCCACCAATCTCGAATATTCGAGAAAGATATT 211546

QY 92 GlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLysGlu 111
DB 211547 CACCGAGATTAGATCGTGTAAACCGAAGACATCTGATATACGACGCTCGTGTGAA 211606

QY 112 ThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThrPro 131
DB 211607 GATGACGAAGTTCAAAATTTTATCTGCTGTGTTTGAAGGAAAAACACACAGCAAGTATT 211666

QY 132 LeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluLeuAlaVal 151
DB 211667 GCGATGATCAATTAATAATGGAGATCAGCGTTCGCAAGATTATGTCACATTAAGATCGT 211726

QY 152 TyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGln 171
DB 211727 TTCGCGCCAGGTCATCGGATTTTACCTATCAGCAAAAGTATGGAATCGGTGATTCGT 211786

QY 172 GlyGlyGlyArgSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuPro 191
DB 211787 GCGGTGGCGGTTCGTACAGCGTGAACAGCGATCGGTTGTCAGGCGCTATTGGC 211846

QY 192 LysLysIleLeuLysLeuLysCysGlyLeuGluLeuLeuSerPheValSerLysValHis 211
DB 211847 AAAAAATATTACCGCAACATTTTGGCAATGAGTGGAGGTTTTTAAAGCAAAATCGT 211906

QY 212 GlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnLeu 231
DB 211907 AATATAAAATTTGCTCCGCGAGAAAGTG-----GGACAAATGATTTGGGAAAGGTAAAC 211960

QY 232 SerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIleAsp 251
DB 211961 AGTAATCCATTTCTTGTCTCATGAAAGTGGCGTAGAAAATTCGATGAATGATCCGT 212020

QY 252 ArgValArgValArgGlyAspSerValGlyGlyValIleThrCysValAlaArgAsnVal 271
DB 212021 GAACTTAAAAAAGAGGAGATCTATTGGCGCAAACTTACTGTATTTCAGAAAAATGTA 212080

QY 272 ProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMet 291
DB 212081 CCTGTAGGATTTGGCGAGCGAGTATTGTACCGTTTAGATGCGCGATCTTCTCACCAGTTA 212140

QY 292 LeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeu 311
DB 212141 ATGGGAATTAATGCAGTAAAGGTGTAGAAATTTGGCGATGGCTTGTCTGTGTTGACNA 212200

QY 312 ThrGlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThrArg 331
DB 212201 CGAGGTTCCGAACATCGTGATGAA-----ATGACACCTTAATGCG-----TTTGAA 212245

QY 332 ThrAsnArgSerGlyValGlnGlyIleSerAsnValGluIleValHisPheLys 351
DB 212246 AGTAATCATGCGCGCGGTATTATTAGCGGAATTTAGTTCAGGACCAACCAATATTCGCCACT 212305

QY 352 ValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGluArg 371
DB 212306 ATTGACCTAAACCAACTTCAAGCATTTACGATCTTGTGCTGCTGATCAATCTTAAATGCT 212365

QY 372 GlnAsnValGluLeuAlaArgGlyArgHisAspProCysValAlaProArgAlaVal 391
DB 212366 GAAGCGGTAGAGTTGTAAACAAAAGTCTGTACAGTCTCTTGTGTGGGATTCGTGCTGTG 212425
```

```
QY 392 ProValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisValAla 411
DB 212426 CCAATTGCGAAGCTATGTCGCGATTGCTTATTAGATCATCTCTTACGTTTAAAGCA 212485

QY 412 GlnCysGlu 414
DB 212486 CAGTGTAAA 212494

RESULT 11
US-09-643-990A-1
; Sequence 1, Application US/09643990A
; Patent No. 6528289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann
; Mark D. Adams
; Owen White
; Hamilton O. Smith
; J. Craig Venter
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville,
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/643,990A
; FILING DATE: 23-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: 1995-06-07
; APPLICATION NUMBER: 08/426,787
; FILING DATE: 1995-04-21
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PBI86P1C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-610-5790
; TELEFAX: 310-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1
```

Alignment Scores:

Pred. No.:	1,12e-80	Length:	1830121
Score:	840.50	Matches:	166
Percent Similarity:	64.74%	Conservative:	69
Best Local Similarity:	45.73%	Mismatches:	121
Query Match:	38.20%	Indels:	7
DB:	4	Gaps:	3

US-10-677-179-8 (1-429) x US-09-643-990A-1 (1-1830121)

```
QY 52 SerGlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGlyGly 71
DB 211427 GCTGGTAATACAACTTGGCAACCTTTCGCTGAGCAACCTTGGAGAGTCACATGGTATT 211486
```

```
QY 72 GlyValGlyCysValIleSerGlyCysProProArgIleProLeuThrGluAlaAspLeu 91
Db 211487 GCATTAGGCTGATCGCGTCGCCACCAAAATCTCGAATTATCCGAGAGATATT 211546
QY 92 GlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLysGlu 111
Db 211547 CAGCCAGATTAGATCGTCGTAACACGAGNAACATCTCGATATACACGCGCTCGTGTGAA 211606
QY 112 ThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrGlyThrProIle 131
Db 211607 GATGACGAAGTTCAAATTTATCTGCTGTTTGAAGGAAAAACACAGGCACAAAGTATT 211666
QY 132 LeuValIleValProAsnThrAspGlnIleGlySerAspHisArgLysIleAlaVal 151
Db 211667 GGGATGATCATTAATAAATCGAGATCAGCGTTCGCAAGATTATGGTGACATTAAGATCGT 211726
QY 152 TyrArgProSerHisAlaAspAlaThrTyrAspPheLysThrGlyValArgAlaValGln 171
Db 211727 TTCCGCCAGGTCATCGGATTTTACCTATACGAAAGATATGGAATCCGTGATTATCGT 211786
QY 172 GlyGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuPro 191
Db 211787 GCGCGTGGCGGTTGTCGACGACGTGAACACAGCGATGCGGTTGCTGCAGGCGCTATTGCG 211846
QY 192 LysLysIleLeuLysLysCysGlyLeuGluIleLeuSerPheValSerLysValHis 211
Db 211847 AAAAAATATTATACGGAACATTTTGGCATTTGAGGTGCGAGGTTTTTAAAGCCAAATCGGT 211906
QY 212 GlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGlu 231
Db 211907 ATATATAAAATTTGCTCCGACGAAAGTG-----GGACAATTTGATTGGGAAAGGTAAC 211960
QY 232 SerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaLeuAsp 251
Db 211961 AGTAATCCATCTTTTGTCTGATGAAAGTGGGTAGAAAAATTCGATGAATGATCGGT 212020
QY 252 ArgValArgValArgGlyAspSerValGlyGlyValIleThrCysValAlaArgAsnVal 271
Db 212021 GAACTTAAAAAAGAGGAGATCTATTGGCGCAAACTTACTGTTATTTCAGAAAAATTGTA 212080
QY 272 ProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMet 291
Db 212081 CCTGTAGGATTGGCGAGCCAGTATTTGACCGTTTAGATGCGCATCTTGTCTCAGCATTA 212140
QY 292 LeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeu 311
Db 212141 ATGGGAATTAATGCAGTAAAAAGGTGAGAAATTTGGCGATGGCTTGTCTGTTGAAACAA 212200
QY 312 ThrGlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThrArg 331
Db 212201 CGAGGTTCCGNAACATCGTATGAA-----ATGACACCTTAATGGC-----TTTGAA 212245
QY 332 ThrAsnArgSerGlyValGlnGlyGlyIleSerAsnValGluIleValHisPheLys 351
Db 212246 AGTAATCATGCGCGCGTATTTTAGCGGAATTTAGTTTCAGGACAACCAATATTCGCCACT 212305
QY 352 ValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGluArg 371
Db 212306 ATTGCACATAAACCAACTTCAAGCATTACGATTCCTGCTGTTCAATCAATCTTAATGGT 212365
QY 372 GlnAsnValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaProArgAlaVal 391
Db 212366 GAAGCCGTAGAGTTGTAAACAAAGGTGCTGACGATCCTTGTGTGGGGATTCTGTCGTGTG 212425
QY 392 ProValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisValAla 411
Db 212426 CCATTTCCGAGCTATGGTGGCGATTGTTTATTATGATCACTCTTACGTTTAAAGGCA 212485
QY 412 GlnCysGlu 414
Db 212486 CAGTGTAAG 212494
```

RESULT 12

```
US-09-248-796A-4099
; Sequence 4099, Application US/09248796A
; Patent NO. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 4099
; LENGTH: 1080
; TYPE: DNA
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (36)
; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unknown
US-09-248-796A-4099
```

```
Alignment Scores:
Pred. No.: 3,25e-84 Length: 1080
Score: 826.50 Matches: 175
Percent Similarity: 66.01% Conservative: 58
Best Local Similarity: 49.58% Mismatches: 103
Query Match: 37.57% Indels: 17
Db: 4 Gaps: 5
```

US-10-677-179-8 (1-429) x US-09-248-796A-4099 (1-1080)

```
QY 68 SerHisGlyGlyGlyValGlyCysValIleSerGlyCysProProArgIleProLeuThr 87
Db 16 TCCCTGGTGCAGATC-----CGNATGTCATAACT 48
QY 88 GluAlaAspLeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSer 107
Db 49 GAAGCGGATATTCAACCACAATTGACTCGTATACGTCAGGTCACCAAGTAAATATCTACC 108
QY 108 ThrArgLysGluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThr 127
Db 109 CCAAGAGATGAAAAAGATTTTATTTGAAATACAAAGTGGAACTGAGATGGGATTAACTTTG 168
QY 128 GlyThrProIleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGlu 147
Db 159 GGTTCACCTATGGATGATGTTGTGAGAATAAAGATCATCGTCTGGTGTATTAGTGAG 228
QY 148 IleAlaAsnValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysThrGlyVal 167
Db 229 ACCGATTTGTTATCCAAAGCAACCAAGTCATGCCACTGGACCTATATACAGAAATACGGAACC 288
QY 168 ArgAlaValGlnGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAla 187
Db 289 AAATCTAGTTCTCTGGTGGTGTAGATCATCTCTAGAGAAACTATTGGTAGAGTTGCCGCA 348
QY 188 GlyAlaLeuProLysLysIleLeuLysLysCysGlyLeuGluIleLeuSerPheVal 207
Db 349 GGAGCCATTGCAGAAAGATTTTGCCAAAGTGAATATGTTGAATTTGTTGCCCTTGTGA 408
QY 208 SerLysValHisGlnValValLeu-----ProGluAspAlaValAsp----- 221
Db 409 AGTGTATTGGTGAATATCTATGAATAAATCTCTCTCAAGATGCAAAATTCACAGAACTT 468
QY 222 TyrGlySerValThrLeuGluGlnIleGluSer---AsnIleValArgCysProAspPro 240
Db 469 TTAACACTATCATCAGAGAACCAAGTCGATGGTGTAGGTCCCATTAAGATGCCAGATGCA 528
QY 241 GluTyrAlaGluLysMetIleAspAlaIleAspArgValArgValArgGlyAspSerVal 260
```

```
Db      529 AATGTCGTGAAGAAATGGTTAAAGTGATTCAGAAATATCGTCATGCTAAAGATTCATT 588
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      261 GlyGlyValIleThrCysValAlaArgAsnValProArgGlyLeuGlySerProValPhe 280
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      589 GGTGGTGTCTGCTACTTGTGTATCAGAAACTGTCCAATTGGATTAGGTGAACCATGTTTC 648
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      281 AspLysLeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPhe 300
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      649 GATAAATTGGAAGCTAAATTTGGCTCATGTCATGTTATCATTTACCCGCTACCAAGGGTTT 708
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      301 GluIleGlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPhe 320
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      709 GAATTTGGTTCTGGATTGAAGGTATCAAGATCCCTGGTTCAAAACATAACGATGCATTC 768
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      321 TyrMetAspLys---AlaGlySerValAlaGlyThrArgThrAsnArgSerGlyGlyValGln 339
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      769 TATTATGATGAAATTTCCGAGAGATTAAGAACTGAAACCAACAATAGTGTGTATCCAA 828
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      340 GlyGlyIleSerAsnValGluIleValHisPheLysValAlaPheLysProThrProSer 359
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      829 GGAGGAATTTCAACCGTGAAACATTTATTTCTCAGTTGCCTTCAATCGGCTGCTACT 888
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      360 IleGlyValLysGlnAsnThrValSerArgGluArgGlnAsnValGluLeuAlaArg 379
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      889 ATCAGTCAAGAACCAAGAACTGCTACTTACGATGTTAAAGTGGTGTGCTGCTAGA 948
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      380 GlyArgHisAspProCysValAlaProArgAlaValProValValGluSerMetAlaAla 399
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      949 GGTAGACAGATCCAGTGTACCCCAAGAGCAGTTCCAAATTTGTAAGCCATGACTGCA 1008
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      400 LeuValLeuMetAspGlnLeuMetAlaHisValAlaGln 412
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1009 TTAGTATTATGTGACGAATATGATTCACAAGCTAGA 1047
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 13
US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; TITLE OF INVENTION: jannaschii
; Patent No. 6503729
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ 'ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
```

```
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103398)..(103398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
```

```

; NAME/KEY: misc feature
; LOCATION: (67435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349491)..(1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1470091)..(1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1569020)..(1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1664854)..(1664855)
; OTHER INFORMATION: n equals a, t, c, or g
; US-08-916-421B-1

```

Alignment Scores: 3.5e-76 Length: 1664976
 Pred. No.: 800.50 Matches: 171
 Percent Similarity: 61.52% Conservative: 64
 Best Local Similarity: 44.76% Mismatches: 122

```

Query Match: 36.39% Indels: 25
DB: 4 Gaps: 6
US-10-677-179-8 (1-429) x US-08-916-421B-1 (1-1664976)
Qy 47 LeuGluVallyValAlaSerGly-----AanThrPheGlyAen 58
Db 357409 TTAATGTAAACTAAGCAATTAATAATTTTGGTGACATTAATGAACACCTATGGGGAT 357468
Qy 59 TyrPheGlnValAlaThrTyGlyGluSerHisGlyGlyValGlyCysValIleSer 78
Db 357469 ATGTTTAGAGTTACAGTTTGGGAGAAAGTCATGGAAGGCTGTGGAGCAGTGTGTCAT 357528
Qy 79 GlyCysProProArgIleProLeuThrGluAlaAspLeuGlnValGlyLeuAspArg 98
Db 357529 GGATGTCAGCTAATCTGCCTTATCTGAAGAGATATCAAAAAGAGCTTGACAGAGA 357588
Qy 99 ArgProGlyGlnSerArgIleThrSerThrArgLysGluThrAspThrCysLysIleLeu 118
Db 357589 AGACGAGGCGAGCATCTCTCAACCAAGAAAGAGAGGATAAAGTTGAATCTTA 357648
Qy 119 SerGlyThrHisGluGlyValThrGlyThrProIleLeuValIleValProAsnThr 138
Db 357649 TCAGGAATTTTGGAGGGGAAAACTACTGGAGCTCTATTTGCTCATAGTCTATAACAAA 357708
Qy 139 AspGlnIleGlySerAspHisArgGluIleAlaAsnValTyArgProSerHisAlaAsp 158
Db 357709 AACATGAGACCTAAAGATTACTCAAAATTAAGAGATACACCAAGACCTGGACATGCGAGAT 357768
Qy 159 AlaThrTyRaspPheLysTyRgLyValArgAlaValGlnGlyGlyArgSerSerGly 178
Db 357769 TTAACCTATAGATTGAAGTATAAAACTATGATTATAGGGGAGGAGGAGGCAAGTGT 357828
Qy 179 ArgLysThrValGlyArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysLeuLys 198
Db 357829 AGAGTAACGATAGGCATGTTATTGGAGGAGCTATTGCTAAAAAGCTTCTATCTTACACA 357888
Qy 199 CysGlyLeuGluIleLeuSerPheValSerLysValHisGlnVal----- 213
Db 357889 TACAACATAAAAAATTATTTGGTTATACCAATAAAGATTGGAAGATTGAAGGAGATTCAGC 357948
Qy 214 -----ValLeuProGluAspAlaValAspTyRgLySerVal-----ThrLeuGluGln 229
Db 357949 TACTATAAAATCCAGAGGTTTTGAAAATCAAAAATCCTTAGAGAGATTAAATAGAGATT 358008
Qy 230 IleGluSerAenIleValArgCysPro-----AspProGluTyRAlaGlyLysMet 246
Db 358009 ATTGAAAGTAATCCATTGAGATGTCCATCAATGAATGAGAAAGAGATGGAGGAGTATGTT 358068
Qy 247 IleAspAlaIleAspArgValArgValArgGlyAspSerValGlyGlyValIleThrCys 266
Db 358069 TTAAGGCAATGGAA-----AATAAGATAGTTGTTGGAGGAGTTGTTGAAATT 358116
Qy 267 ValAlaArgAenValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGlu 286
Db 358117 GTTCATTAAATGTTCTCTGTTGGAGTTGGAATCCAATATTCAATAAGTTAAATGGAGAA 358176
Qy 287 LeuAlaLysAlaMetLeuSerIleProAlaSerAenGlyPheGluIleGlySerGlyPhe 306
Db 358177 TTGGCAAGAGCTTAAATGAGTATAAATGCTGTTAAAGAGAGTTGAGATGGGCGCTGGTTTT 358236
Qy 307 AlaGlyThrAspLeuThrGlySerGluHisAenAspGluPheTyRMetAspLysAlaGly 326
Db 358237 AAAGCGCTGAGATGTATGGAAGTGAGATGAACATGAGATGATTTTGTATGACGACACAA 358296
Qy 327 SerValArgThrArgThrAsnArgSerGlyGlyValGlnGlyGlyIleSerAsnValGlu 346
Db 358297 AATATAAGATTCAAAACAAACAACTGCGGTGGCATATTGGGAGGAGATTAGCTGTGGAAT 358356
Qy 347 IleValHisPheLysValAlaPheLysProThrProSerIleGlyValLysGluAsnThr 366
Db 358357 CCAATAGTTTAAAGATTGCAAGTAAGCAACACTTCAATAGGTAAAGCAAAAAACC 358416

```


QY 367 ValSer---ArgGluArgGlnAenValGluLeuAlaArgGlyArgHisAspProCys 385
Db 358417 ATAAATTAAACCTTAGAAATGTTGAAATTTGAAGTGAAGGACACGACCCAGTT 358476
QY 386 ValAlaProArgAlaValProValValGluSerMetAlaLeuValLeuMetAspGln 405
Db 358477 ATAGTTCCAGGATGTTCCAGTGGCTGAAGCAATGCTGTATATACCTTAGCTGATTTG 358536
QY 406 LeuMet 407
Db 358537 ATGATT 358542
RESULT 14
US-09-692-570-1
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98256)..(98256)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g

FEATURE:		Best Local Similarity:	44.76%	Mismatches:	122
NAME/KEY: misc feature		Query Match:	36.39%	Indels:	25
LOCATION: (657081)..(657081)		DB:	4	Gaps:	6
FEATURE:		US-10-677-179-8 (1-429) x US-09-692-570-1 (1-1664976)			
NAME/KEY: misc feature		Qy	47	LeuGluVallylsAlaSerGly-----AenThrPheGlyAsn	58
LOCATION: (674435)..(674435)		Db	357409	TTAAATGTAAACCTAAGCAATTTTGGTGACATTATGAACACCTATGGGGAT	357468
FEATURE:		Qy	59	TyrPheGlnValAlaThrTyGlyGluSerHisGlyGlyValGlyCysValIleSer	78
LOCATION: (657203)..(657203)		Db	357469	ATGTTTAGAGTTACAGTTTTTGGAGAAAGTCATCGAAAGCGTGTGGAGCAGTTGTTCAT	357528
FEATURE:		Qy	79	GlyCysProProArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArg	98
LOCATION: (682442)..(682442)		Db	357529	GGATGTCAGCTAATCTGCCTTTATCTGAAGAGATATCCAAAAGAGCTTGACAGAGA	357588
FEATURE:		Qy	99	ArgProGlyGlnSerArgIleThrSerThrArgLysGluThrAspThrCysLysIleLeu	118
LOCATION: (713652)..(713652)		Db	357589	AGACCAGGCAGACATCTTCTCAACCAAGAAAGAGAGGATAAGTTGAANTCTTA	357648
FEATURE:		Qy	119	SerGlyThrHisGluGlyValThrThrGlyThrProIleLeuValIleValProAsnThr	138
LOCATION: (741684)..(741684)		Db	357649	TCAGGAATTTTGGAGGGAAAACTACTGGAGCTCTATTGCTCAATAGTCTATAACAA	357708
FEATURE:		Qy	139	AspGlnIleGlySerAspHisArgGluIleAlaAsnValTyArgProSerHisAlaAsp	158
LOCATION: (779455)..(779455)		Db	357709	AACATGAGACCTAAAGATTACTCAAAAATTAAGATACACCAAGACCTGGACATGCAGAT	357768
FEATURE:		Qy	159	AlaThrTyraaspPheLysTyGlyValArgAlaValGlnGlyGlyArgSerSerGly	178
LOCATION: (779676)..(779676)		Db	357769	TTAACCTATAGATTGAAGTATAAAACTATGATTATAGGGAGGAGGAGGCAAGTGGT	357828
FEATURE:		Qy	179	ArgLysThrValGlyArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysLeuLys	198
LOCATION: (855539)..(855539)		Db	357829	AGAGTAACGATAGGCATGTTATTGGAGGAGCTATTGCTAAAAAGCTTCTATCTTACACA	357888
FEATURE:		Qy	199	CysGlyLeuGluIleLeuSerPheValSerLysValHisGlnVal-----	213
LOCATION: (871619)..(871619)		Db	357889	TACAACATAAAATTTATTGGTTATACCATAAAGATTGGAAGATTGAAGGAGATTTCAGC	357948
FEATURE:		Qy	214	-----ValLeuProGluAspAlaValAspTyryGlySerVal-----ThrLeuGluGln	229
LOCATION: (1084830)..(1084830)		Db	357949	TACTATAAAATCCAGAGGTTTTTGAAATCAAAATCCITAGAGAGATTAAATAGAGATT	358008
FEATURE:		Qy	230	IleGluSerAenIleValArgCysPro-----AspProGluTyryAlaGluLysMet	246
LOCATION: (1096846)..(1096846)		Db	358009	ATTGAAAGTAAATCCATTGAGATGTCCATCAATGAATGAGAAAGAGATGGAGGAGTATGTT	358068
FEATURE:		Qy	247	IleAspAlaIleAspArgValArgValArgGlyAspSerValGlyGlyValIleThrCys	266
LOCATION: (1119881)..(1119881)		Db	358069	TTAAAGGCAATGGAA-----AATAAGATAGTGTGGAGGAGTGTGGAAATT	358116
FEATURE:		Qy	267	ValAlaArgAenValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGlu	286
LOCATION: (1130881)..(1130881)		Db	358117	GTTGCATTAAATGTTCTCTGTGGAGTTGGAATCCATATTCATATAAGTTAAATGGAGAA	358176
FEATURE:		Qy	287	LeuAlaLysAlaMetLeuSerIleProAlaSerAenGlyPheGluIleGlySerGlyPhe	306
LOCATION: (1310988)..(1310988)		Db	358177	TTGGCAAGAGCTTTAATGAGTATAAATCTGTTAAAGAGTGTGAGATAGGGCGTGGTTTT	358236
FEATURE:		Qy	307	AlaGlyThrAspLeuThrGlySerGluHiAsnAspGluPheTyryMetAspLysAlaGly	326
LOCATION: (1313224)..(1313224)		Db	358237	AAAGCGCTGAGATGTATGGAAAGTGAGATGAACCATGAGATGTTTGTATGATGACACAAA	358296
FEATURE:		Qy	327	SerValArgThrArgThrAsnArgSerGlyGlyValGlnGlnGlyIleSerAsnValGlu	346
LOCATION: (1349473)..(1349473)		Db	358297	AATATAGATTCAAAACAAACAACTCGCGTGGCATATTGGGAGGAAATTAGCTGTGAACT	358356
FEATURE:		Qy	347	IleValHisPheLysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThr	366
LOCATION: (1664976)..(1664976)		Db	358357	CCAATAGTTTTTAAGAAATTGCAAGTAAAGCCACACCTTTCAATAGGTAAAGCAAAAAACC	358416

Alignment Scores:

Pred. No.: 3.5e-76

Score: 800.50

Percent Similarity: 61.52%

Length:

Matches:

Conservative: 64

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 27, 2005, 04:44:29 ; Search time 727 Seconds
(without alignments)
3861.053 Million cell updates/sec

Title: US-10-677-179-8

Perfect score: 2200

Sequence: 1 MTTVPKQQVAHSRARLAPR.....VAQCEMFALTALQEPVGSF 429

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 7331713 seqs, 327154945 residues

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp

-Q=/cgn2_1/USPTO_spool/p/US10677179/runat_25082005_150241_24891/app_query.fasta_1.583

-DB=Published Applications NA -QWMT=fastap -SUFFIX=rnpb -MINMATCH=0.1

-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62

-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100

-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0

-MAXLEN=2000000000 -USER=US10677179@cgn 1 1 723 @runat_25082005_150241_24891

-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100

-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5

-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:
- 22: /cgn2_6/ptodata/2/pubpna/US10J_NEW_PUB.seq:
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2200	100.0	1626	18	US-10-677-179-7
2	2090	95.0	1845	20	US-10-425-115-138337
3	2016.5	91.7	1496	18	US-10-425-114-30604
4	2011.5	91.4	1541	20	US-10-425-115-138335
5	1800.5	81.8	1792	20	US-10-425-115-171490
6	1797.5	81.7	1635	18	US-10-677-179-1
7	1790.5	81.4	1603	18	US-10-425-114-32244
8	1724	78.4	1353	19	US-10-437-963-99207
9	1512	68.7	1674	18	US-10-425-114-8982
10	1512	68.7	2094	18	US-10-424-599-1811
11	1454	66.1	1311	19	US-10-696-616-27
12	1257	57.1	966	18	US-10-677-179-11
13	1191.5	54.2	1089	17	US-10-369-493-26364
14	1186.5	53.9	1086	17	US-10-369-493-42555
15	1182.5	53.8	1086	17	US-10-369-493-43897
16	1066	48.5	1015	18	US-10-677-179-5
17	1035	47.0	1080	17	US-10-369-493-44740
18	981	44.6	1068	17	US-10-369-493-34106
19	938	42.6	1131	17	US-10-369-493-45609
20	934	42.5	1137	21	US-10-741-849-6185
21	898	40.8	1292	18	US-10-424-599-1812
22	888.5	40.4	1074	17	US-10-282-122A-12647
23	869.5	39.5	1421	17	US-10-369-493-27659
24	863	39.2	724	20	US-10-425-115-171487
25	862	39.2	1086	17	US-10-282-122A-42063
26	860	39.1	1086	9	US-09-815-242-6135
27	860	39.1	1086	17	US-10-369-493-47170
28	860	39.1	1086	17	US-10-282-122A-20439
29	860	39.1	3100	21	US-10-489-273-5
30	858	39.0	1056	17	US-10-282-122A-24843
31	857	39.0	1086	9	US-09-815-242-7611
32	855	38.9	1083	17	US-10-282-122A-23613
33	850	38.6	1083	17	US-10-282-122A-32881
34	846.5	38.5	1074	17	US-10-282-122A-30794
35	845	38.4	640681	9	US-09-790-988-1
36	843	38.3	1080	17	US-10-369-493-44842
37	840.5	38.2	1071	14	US-10-260-877-127
38	840.5	38.2	1074	9	US-09-815-242-6884
39	840.5	38.2	1074	17	US-10-282-122A-21915
40	840.5	38.2	1077	17	US-10-369-493-44300
41	840.5	38.2	1830121	17	US-10-329-670-1
42	840.5	38.2	1830121	20	US-10-158-865-1
43	840.5	38.2	1830121	22	US-10-981-687-1
44	840	38.2	1086	17	US-10-282-122A-41210
45	830.5	37.8	726	19	US-10-767-701-9032

ALIGNMENTS

RESULT 1

US-10-677-179-7

Sequence 7, Appli US/10677179

Publication No. US20040082050A1

GENERAL INFORMATION:

APPLICANT: Cahoon, Rebecca E.

APPLICANT: Falco, Saverio Carl

APPLICANT: Pember, Stephen O.

TITLE OF INVENTION: Chorismate Biosynthesis Enzymes

FILE REFERENCE: BB-1159-A

CURRENT APPLICATION NUMBER: US/10/677,179

CURRENT FILING DATE: 2003-10-02

PRIOR APPLICATION NUMBER: US/09/743,207

PRIOR FILING DATE: 2001-01-04

PRIOR APPLICATION NUMBER: PCT/US99/16353

PRIOR FILING DATE: 1999-07-20

! PRIOR APPLICATION NUMBER: 60/093,611
! PRIOR FILING DATE: 1998-07-21
! NUMBER OF SEQ ID NOS: 14
! SOFTWARE: Microsoft Office 97
! SEQ ID NO 7
! LENGTH: 1626
! TYPE: DNA
! ORGANISM: Zea mays
US-10-677-179-7

Alignment Scores:
Pred. No.: 2,91e-239 Length: 1626
Score: 2200.00 Matches: 429
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-677-179-8 (1-429) x US-10-677-179-7 (1-1626)

Qy 1 MetThrThrValProLysProGlnValAlaHisSerArgAlaArgLeuAlaProArg 20
Db 79 ATGACGACCGTGCACCAAGCCACAGCAGGTGGCGCACTCACGGGCACGGCTCGCACCCCGC 138
Qy 21 AlaileGlyAlaLeuLeuGluPheAlaProAlaSerSerLeuArgPheAlaValHis 40
Db 139 GCGATCGGGCGCTTGCTGGAGTTTGCCCCACAGCTCTCTCTCCCTCCGCTTCGCGTGGAC 198
Qy 41 ArgCysArgThrAlaArgLeuGluValLysAlaSerGlyAsnThrPheGlyAsnThrPhe 60
Db 199 CGCTGCCGCACTGCTCGCTAGAGGTGAAGCATCTCGAAACACAGTTTGGAAACTACTTTT 258
Qy 61 GlnValAlaThrThrGlyLysSerHisGlyGlyValGlyCysValLysSerGlyCys 80
Db 259 CAGGTTCGAACCTATGGTGAATCTCATGGGGTGGTGTGGTGTGTATCAGTGGTGTGT 318
Qy 81 ProProArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArgArgPro 100
Db 319 CCACCTAGAAATCCACTCACTGAGGCGACACTCAAGTTGAATCGATCGAAGACGGCCC 378
Qy 101 GlyGlnSerArgIleThrSerThrArgLysGluThrAspThrCysLysIleLeuSerGly 120
Db 379 GGACAGACAGAAATCACTCCACAGGAAGGAGACTGATACATGCAAAATCTCTGTCAGG 438
Qy 121 ThrHisGluGlyValThrGlyThrProIleValIleValProAsnThrAspGln 140
Db 439 ACACATGAAGGGGTGACTACTGGAAACGCCAATCTTGTATTGTCTCCCAACACAGATCAA 498
Qy 141 IleGlySerAspHisArgGluIleAlaAsnValTyrArgProSerHisAlaAspAlaThr 160
Db 499 ATAGGCAGTGATCACCGTGAATAGCCAAATGTGTACCGACCTTCTCATGCGAGCGCACT 558
Qy 161 TyrAspPheLysTyrGlyValArgAlaValGlnGlyGlyArgSerSerGlyArgLys 180
Db 559 TATGACTTCAAGTACGGTGTAGAGCTGTACAGGGAGGTGGAGGTCTCTCGGCAGAAAA 618
Qy 181 ThrValGlyArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysLeuLysCysGly 200
Db 619 ACCGTGGAAGGGTGGCTGCGAGGGGCCCTCCCAAGAAAAATCTTAAAGCTCAAAATGTGGA 678
Qy 201 LeuGluIleLeuSerPheValSerLysValHisGlnValValLeuProGluAspAlaVal 220
Db 679 TTAGAGATCTTGCTGTTGTTTCCAAAGTGATCAGGTTGTCTCCCAAGAGACGGGTT 738
Qy 221 AspTyrGlySerValThrLeuGluIleGluSerAsnIleValArgCysProAspPro 240
Db 739 GATTATGGGTCTGTAACTTGGAAACAGATAGAGCAACATCGTTAGATGCTCTGATCCA 798
Qy 241 GluTyrAlaGluLysMetIleAspAlaIleAspArgValArgValArgGlyAspSerVal 260
Db 799 GAGTACGCAGAGAAAGATATAGACCAATCAGACAGTACAGGTTCGAGGGGGATTCGGTTC 858
Qy 261 GlyGlyValIleThrCysValAlaArgAsnValProArgGlyLeuGlySerProValPhe 280

Db 859 GGTGGAGTGATCATCATCGCTCGTAGAAAGCTTCTCGCGGGCTCGGTTCTCTGTGTTTC 918
Qy 281 AspLysLeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPhe 300
Db 919 GACAAGCTCGAATCCGAATCGCAAAAGCTATGCTTTCTATTCTCTGCGAGCAACGGGTTTC 978
Qy 301 GluIleGlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPhe 320
Db 979 GAGATTGGCAGCGGATTTCGCCGGACCGACTTGACAGAAAGTGAGCATAAATGATGAGTTT 1038
Qy 321 TyrMetAspLysAlaGlySerValArgThrArgThrAsnArgSerGlyGlyValGlnGly 340
Db 1039 TATATGATTAAGCTGGAAGTGTCAAGACACGAGCACTAATCGCTCGGGTGTGTGACAGGA 1098
Qy 341 GlyIleSerAsnValGluIleValHisPheLysValAlaPheLysProThrProSerIle 360
Db 1099 GGGATATCGAATGTTGAGATTGTGCACTTCAAAAGTTGCTTTAAAGCCGACACCATCTATC 1158
Qy 361 GlyValLysGlnAsnThrValSerArgGluArgGlnAsnValGluLeuLeuAlaArgGly 380
Db 1159 GGGGTGAAAACAGAACACCGGTGTCAAGGGAGCGTCAGAACGTTGAGCTTCTAGCAAGAGGG 1218
Qy 381 ArgHisAspProCysValAlaProArgAlaValProValValGluSerMetAlaAlaLeu 400
Db 1219 CGCATGACCCATCGCTCGCCCTCGAGCTGTTCTGTGGTGGATCCATGCGCGCTTG 1278
Qy 401 ValLeuMetAspGlnLeuMetAlaHisValAlaGlnCysGluMetPheAlaLeuAsnThr 420
Db 1279 GTCTCATGACACGAGCTGATGGCGCAGCTGGCTCAGTGCAGATGTTTCGGCTCAATACT 1338
Qy 421 AlaLeuGlnGluProValGlySerPhe 429
Db 1339 GCACTTCAAGAACAGTTGGCTCTTTTC 1365

RESULT 2

US-10-425-115-138337
! Sequence 138337, Application US/10425115
! Publication NO. US20040214272A1
! GENERAL INFORMATION:
! APPLICANT: La Rosa, Thomas J.
! APPLICANT: Kovalic, David K.
! APPLICANT: Zhou, Yihua
! APPLICANT: Cao, Yongwei
! TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
! FILE OF INVENTION: Plants
! FILE REFERENCE: 38-21(53222)B
! CURRENT APPLICATION NUMBER: US/10/425,115
! CURRENT FILING DATE: 2003-04-28
! NUMBER OF SEQ ID NOS: 369326
! SEQ ID NO 138337
! LENGTH: 1845
! TYPE: DNA
! ORGANISM: Zea mays
! FEATURE:
! OTHER INFORMATION: Clone ID: MRT4577_57640C.1
US-10-425-115-138337

Alignment Scores:
Pred. No.: 1,09e-226 Length: 1845
Score: 2090.00 Matches: 422
Percent Similarity: 84.80% Conservative: 2
Best Local Similarity: 84.40% Mismatches: 5
Query Match: 95.00% Indels: 71
DB: 20 Gaps: 2

US-10-677-179-8 (1-429) x US-10-425-115-138337 (1-1845)

Qy 1 MetThrThrValProLysProGlnValAlaHisSerArgAlaArgLeuAlaProArg 20
Db 111 ATGACGACCGTGCACCAAGCCACAGCAGGTGGCGCACTCACGGGCACGGCTCGCACCCCGC 170
Qy 21 AlaileGlyAlaLeuLeuGluPheAlaProAlaSerSerSerLeuArgPheAlaValHis 40

Db 171 GCGATCGCGCCCTTCTGAGTTTGGCCGAGCTCTCTCTCCCTCGCTTCGCGTGCAC 230
Qy 41 ArgCysArgThrAlaArgLeuGluValLysAlaSerGlyAsnThrPheGlyAsnTyrPhe 60
Db 231 CGCTGCGGACCTGCTCGCTAGAGGTGAAGGATCTCGGAACACAGTTTGGAACTACTTT 290
Qy 61 GlnValAlaThrTyrGlyGluSerHisGlyGlyGlyValGlyCysValIleSerGlyCys 80
Db 291 CAGGTTGCAACCTATGGTGAATCTCATGGGGGTGGTGGTGGTGGTGGTGGTGGTGGT 350
Qy 81 ProProArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArgArgPro 100
Db 351 CCACCTAGAAATTCCTACTGAGGCGAGACCTACAAGTTGAATCGATCGAAGACGGGCC 410
Qy 101 GlyGlnSerArgIleThrSerThrArgLysGluThrAspThrCysValIleLeuSerGly 120
Db 411 GGTGAGAGAGAATAACCTCCCAAGAAAGAGAGCTGATACATGCAAAATTCCTGTCAGGG 470
Qy 121 ThrHisGluGlyValThrThrGlyThrProIleLeuValIleValProAsnThrAspGln 140
Db 471 ACACATGAAGGGGTGACTACTGGTACGCCAATCTTGTATTGTCTCCAAACACAGATCA 530
Qy 141 IleGlySerAspHisArgGluIleAlaAsnValTyrArgProSerHisAlaAspAlaThr 160
Db 531 ATAGGCACTGATCACCGTGAAATAGCAATGTGTACCGACCTTCTCATGCGACGCAACT 590
Qy 161 TyrAspPheLysTyrGlyValArgAlaValGlnGlyGlyArgSerSerGlyArgLys 180
Db 591 TATGACTTCAAGTACGGTGTATAGAGCTGTACAGGAGGTGGAGGTCTGCGGCGAGAGA 650
Qy 181 ThrValGlyArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysLeuLysCysGly 200
Db 651 ACCGTTGAGGTGGCTGCGAGGGCCCTCGCAAGAAAATTTCTTAAGCTCAAAATGCGA 710
Qy 201 LeuGluIleLeuSerPheValSerLysValHisGlnValValLeuProGluAspAlaVal 220
Db 711 TTAGAGATCTTGTCTGTTGTTTCCAAAGTGATCAGGTTGTCTCCAGAGAGACGGGT 770
Qy 221 AspTyrGlySerValThrLeuGluGlnIleGluSerAsnIleValArgCysProAspPro 240
Db 771 GATTATGGTCTGTAACTTGGAAATATAGAGACCAATCATGTTAGATGCTCTGATCCA 830
Qy 241 GluTyrAlaGluLysMetIleAspAlaIleAspArgValArgValArgLysAspSerVal 260
Db 831 GAGTACGACAGAGAATGATAGACCAATCGACAGATGAGGTTTCGAGGGGATTCGGTC 890
Qy 261 GlyGlyValIleThrCysValAlaArgAsnValProArgGlyLeuGlySerProValPhe 280
Db 891 GGTGAGTGTATCATGCTGCTAGAAACGTTCTCTCGGGGCTCGGTTCTCTGTTGTC 950
Qy 281 AspLysLeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPhe 300
Db 951 GACAGCTCGAATCGAATCGCAATCGCAAAAGCTATGCTTCTATTCTCGCAGCAACGGGTC 1010
Qy 301 GluIleGlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPhe 320
Db 1011 GADATGCGCGGATTCGCGGGACCGACTTGACAGGAAGTGACATAATGATGAGTTT 1070
Qy 321 TyrMetAspLysAlaGlySerValArgThrArgThrAsnArgSerGlyGly 337
Db 1071 TATATGATAGGCTGGAAGTGTGAGACACGAGCTAATCGCTCGGGGGTGTGAGGTT 1130
Qy 337 ----- 337
Db 1131 TTTTTCCTTCATCTTAATAAACCCCTATCGATTGTGACTGTGCAATGTTCTCTAAAA 1190
Qy 338 -----ValGlnGlyGlyIleSerAsnValGluIleVal 348
Db 1191 AAATGGAAATCGTTTGTGTTGTAATGTGCGAGGGAGGATATCGAATGGTGGAGATTGTG 1250
Qy 349 HisPheLysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSer 368

Db 1251 CACTTCAAAGTTGCTTTTAAAGCGACACCATCTATCGGGGTGAAACAGAACACACTGTGTCA 1310
Qy 369 ArgGluArgGlnAsnValGluLeuValGluLeuAlaArgGlyArgHisAspProCysValAlaPro 388
Db 1311 AGGAGAGCTGAGAACGTTGAGCTTTGTCGAAGAGGGGCCCATGACCCATCGCTCGCCCT 1370
Qy 389 Arg ----- 389
Db 1371 CGAGGTAATGCGACCGTGTATGAGCTCCTCGCTGTAGTGTCTGTCTACCATCTTTT 1430
Qy 390 -----Al 390
Db 1431 ATTTCTTCCCGCGACGCTGCGATACGAGGTTTGTCTCTGACGTGGCTTCTGCGAGC 1490
Qy 390 aValProValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisVa 410
Db 1491 TGTCTCTGTGTGGAATCCATGGCGCGCTTGGTCTCTGTGACCGACGCTGTATGGCGCAGT 1550
Qy 410 lAlaGlnCysGluMetPheAlaLeuAsnThrAlaLeuGlnGluProValGlySerPhe 429
Db 1551 GGGCCAGTGGCAGATGTTGGCGCTCAATGTGCACTTCAAGAACAGTGGCTCTTTC 1608

RESULT 3

US-10-425-114-30604
; Sequence 30604, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 30604
; LENGTH: 1496
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73045G02_FLI
US-10-425-114-30604

Alignment Scores:
Pred. No.: 1,82e-218 Length: 1496
Score: 2016.50 Matches: 398
Percent Similarity: 93.01% Conservatives: 1
Best Local Similarity: 92.77% Mismatches: 3
Query Match: 91.66% Indels: 27
DB: 18 Gaps: 1

US-10-677-179-8 (1-429) x US-10-425-114-30604 (1-1496)

Qy 1 MetThrThrValProLysProGlnValAlaHisSerArgAlaArgLeuAlaProArg 20
Db 68 ATGACGACCGTGCACCAAGCCACACGAGGTGGCGCACTCAGGGCAGCGCTCGCACCCGCG 127
Qy 21 AlaIleGlyAlaLeuLeuGluPheAlaProAlaSerSerLeuArgPheAlaValHis 40
Db 128 GCGATCGGCGCTTGTGAGTGTGGAGTTTGGCCCGACGCTCTCTCCCTCGCGTGGCGAC 187
Qy 41 ArgCysArgThrAlaArgLeuGluValLysAlaSerGlyAsnThrPheGlyAsnTyrPhe 60
Db 188 CGCTGCGCACTGCTCGCTAGAGTCAAGGCATCTCGAAACACAGTTTGGAACTACTTT 247
Qy 61 GlnValAlaThrTyrGlyGluSerHisGlyGlyValGlyCysValIleSerGlyCys 80
Db 248 CAGGTTGCAACCTATTGTGAATCTCATGGGGGTGGTGTGGTGTGTGTATCAGTGGTGT 307

668	Db		TTAGAG-----	673
221	Qy	AspTyrGlySerValThrLeuGluGlnIleGluSerAsnIleValArgCysProAspPro	240	
674	Db	-----	-----ATAGAGAGCAACATCGTTAGATGTCCTGATCCA	706
241	Qy	GluTyrAlaGluLysMetIleAspAlaIleAspArgValArgValArgGlyAspSerVal	260	
707	Db	GAGTACGCAGAGAAGATGATAGACGCAATCGACAGAGTACAGATTCGAGGGGATTCGGTC	766	
261	Qy	GlyGlyValIleThrCysValAlaArgAsnValProArgGlyLeuGlySerProValPhe	280	
767	Db	GGTGGAGTGATCACATCGTCGCTAGAAACGTTCTCGCGGCGTCGGTTCCTCTGTGTTC	826	
281	Qy	AspLysLeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSerAsnGlyPhe	300	
827	Db	GACAAGCTCGAATCGGAATCGCAAAAGCTATGCTTCTATTCTCGAGCAACCGGTTTC	886	
301	Qy	GluIleGlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAsnAspGluPhe	320	
887	Db	GAGATTGGCAGCGGATTTCGCGGAGCCGACTTGACAGAAAGTGACGATAATGATGAGTTT	946	
321	Qy	TyrMetAspLysAlaGlySerValArgThrArgThrAsnArgSerGlyValGlnGly	340	
947	Db	TATATGATTAAGCGTGAAGTGTGAGACACGGACTAATCGCTCGGTGTGTGTCAGGGA	1006	
341	Qy	GlyIleSerAsnValGluIleValHisPheLysValAlaPheLysProThrProSerIle	360	
1007	Db	GGGATATCGAATGCTGAGATTGTGCACCTTCAAAGTTCCTTTAAGCGCACCATCTATC	1066	
361	Qy	GlyValLysGlnAsnThrValSerArgGluArgGlnAsnValGluLeuLeuAlaArgGly	380	
1067	Db	GGGGTGAACAGAAACACCGTGTCAAGGGAGCGTCAGAAAGTTCGAGCTTCGCGCAAGAGG	1126	
381	Qy	ArgHisAspProCysValAlaProArgAlaValProValValGluSerMetAlaAlaLeu	400	
1127	Db	CGCCATGACCCATCGTCGCGCCCTCGAGCTGTTCTCTGTGTGAATCCATGGCGCGGTG	1186	
401	Qy	ValLeuMetAspGlnLeuMetAlaHisValAlaGlnCysGluMetPheAlaLeuAsnThr	420	
1187	Db	GTGCTCATGACCAGCTGATGGCGCAGCTGGCCAGTGCAGAGATGTTTCGGCTCAATGCT	1246	
421	Qy	AlaLeuGlnGluProValGlySerPhe	429	
1247	Db	GCATTCAGAACCAGTTGGCTCTTTC	1273	

RESULT 5

```

US-10-425-115-171490
; Sequence 171490, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 171490
; LENGTH: 1792
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MFT4577_87981C.1
US-10-425-115-171490

```

Alignment Scores:

Alignment scores:		
Pred. No.:	8.17e-194	Length: 1792
Score:	1800.50	Matches: 350

Db 1132 GATGAGGCTGGAATATGTGAGCAGCAACTAATCGCTCAGCGGTGTTTCAGGAGGGATA 1191
Qy 343 SerAenValGluIleValHisPheLysValAlaPheLysProThrProSerIleGlyVal 362
Db 1192 TCAATGTGTAATATTATTCTTCAAGTGGCTTTTAAGCCACAGCAACTATCGGAAG 1251
Qy 363 LysGlnAenThrValSerArgGluArgGlnAenValGluLeuLeuAlaAatGlyArgHis 382
Db 1252 AAGCAAAATACGTGTCTAAGGAGCATGAGATGTTGAACCTTTTGGCAAGGGGGCGCAT 1311
Qy 383 AspProCysValAlaProArgAlaValProValValGluSerMetAlaAlaLeuValLeu 402
Db 1312 GACCCCTGTGTGTCCTCGAGCTGTTCTATGTTGGAATCCATGGCTGCGCTGCTG 1371
Qy 403 MetAenGlnLeuMetAlaHisValAlaGlnCysGluMetPheAlaLeuAenThrAlaLeu 422
Db 1372 ATGGACCAAGCTCATGGCGCATATTGCCAGTGTGAGATGTTTCGGCTGAACCTTGGCCCTA 1431
Qy 423 GlnGluProValGlySer 428
Db 1432 CAAGAGCCCATTTGGCTCT 1449

RESULT 6

US-10-677-179-1
; Sequence 1, Application US/10677179
; Publication No. US20040082050A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, Steven O.
; APPLICANT: Pember, Stephen O.
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-A
; CURRENT APPLICATION NUMBER: US/10/677,179
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: US/09/743,207
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: PCT/US99/16353
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1635
; TYPE: DNA
; ORGANISM: Zea mays
US-10-677-179-1

Alignment Scores:
Pred. No.: 1,578-193 Length: 1635
Score: 1797.50 Matches: 349
Percent Similarity: 89.44% Conservative: 32
Best Local Similarity: 81.92% Mismatches: 42
Query Match: 81.70% Indels: 3
Gaps: 2

US-10-677-179-8 (1-429) x US-10-677-179-1 (1-1635)

Qy 4 ValProLysProGlnGlnValAlaHisSerArgAlaArgLeuAlaProArgAlaIleGly 23
Db 64 GTGTGCGAGCGCGGTGTCGCCAGGGCGTCCACAGCGTTTCTCCCGCGGGATAGGC 123
Qy 24 AlaLeuLeuGluPheAlaProAlaSerSerSerLeuArgPheAlaVal---HisArgCys 42
Db 124 GCGCTCCGGAGTCGCGCC---ACGTCCCTCCGGTTATCCGTCGGCGCGCGTGC 177
Qy 43 ArgThrAlaArgLeuGluValLysAlaSerGlyAenThrPheGlyAenThrPheGlnVal 62
Db 178 CGCGCTCCAGCTAGAGTGAAGCATCAGGAATGTGTTCCGGAATCTTCCAGTT 237
Qy 63 AlaThrTyrGlyGluSerHisGlyGlyValGlyCysValIleSerGlyCysProPro 82
Db 238 GCAACCTATGCGAATCCATGGAGGGGTGTTGGTTTCGGTTATCAGTGGCTGCCACCC 297

Qy 83 ArgIleProLeuThrGluAlaAspLeuGlnValGluLeuAspArgArgProGlyGln 102
Db 298 AGAATTCCTCTCACTAGGCGAGCATCGAAGTAGAACTCGATAGAAAGCGTCCGGGTCAA 357
Qy 103 SerArgIleThrSerThrArgLysGluThrAspThrCysLysIleLeuSerGlyThrHis 122
Db 358 AGTAGAATTAACAACCCAGAAAGGAGACTGATACATGCAAAATTTCTATCAGGACACAT 417
Qy 123 GluGlyValThrThrGlyThrProIleLeuValIleValProAenThrAspGlnIleGly 142
Db 418 GATGGATGACTACTGTGTACCAATTCACGTCTTTGTGCCAAACACAGATCAAGGGGT 477
Qy 143 SerAspHisArgGluIleAlaAenValTyrArgProSerHisAlaAspAlaThrTyrAsp 162
Db 478 GGTGATTTACAGTAGTAATGCTTAAGGGGTACAGACCATCCCATCAGATGCAACTATGAC 537
Qy 163 PheLysTyrGlyValArgAlaValGlnGlyGlyArgSerSerGlyArgLysThrVal 182
Db 538 TTCAAGTATGGAGTTAGAGCTGTGAGGGAGGTGGAAAGTTCATGCCAGAGAAACCAT 597
Qy 183 GlyArgValAlaAlaGlyAlaLeuProLysLysIleLeuLysLysCysGlyLeuGlu 202
Db 598 GGCAGGGTGGCTGCAGGAGCTTTGCCAAGAAATAATTCTAAAGCTCAATCAGGAGTGGAG 657
Qy 203 IleLeuSerPheValSerLysValHisGlnValValLeuProGluAspAlaValAspTyr 222
Db 658 ATCTTGGCATTTGTTTCTAAAGTGCACCAAGTCGTACTTCCAGAAGATGCGATGATAT 717
Qy 223 GlySerValThrLeuGluGlnIleGluSerAsnIleValArgCysProAspProGluTyr 242
Db 718 GAGACTGTAACTTGGNACATATAGAGAGCAACATCGTTAGATGTCCTGATCCAGATAT 777
Qy 243 AlaGluLysMetIleAspAlaIleAspArgValArgValArgGlyAspSerValGlyGly 262
Db 778 GCAGAGAAGATGATGCTGCCATTGATCGGTACGAGTGTAGAGAGATTCATTTGGTGGG 837
Qy 263 ValIleThrCysValAlaArgAenValProArgGlyLeuGlySerProValPheAspLys 282
Db 838 GTCGTACATGCATTCGAAAGAAATGTCCTCGTGGTCTTGGCTCTCTCTGTTTTGACAA 897
Qy 283 LeuGluSerGluLeuAlaLysAlaMetLeuSerIleProAlaSerAenGlyPheGluIle 302
Db 898 CTTGAAGCTGAATGCGCAAAAGCCATGCTTCTTCTTCTCCAGCAAGGGGTTTGAGATT 957
Qy 303 GlySerGlyPheAlaGlyThrAspLeuThrGlySerGluHisAenAspGluPheTyrMet 322
Db 958 GGCAGTGGGTTCGCTGCTACCGACTTTTACTGGAAGTGAGCATTAATGATGATTCATATG 1017
Qy 323 AspLysAlaGlySerValArgThrArgThrAsnArgSerGlyGlyValGlnGlyGlyLe 342
Db 1018 GATGAGGCTGGAATGTGAGGACACGAACCTAAATCGCTCAGGGGTGTTTCAGGGAGGATA 1077
Qy 343 SerAenValGluIleValHisPheLysValAlaPheLysProThrProSerIleGlyVal 362
Db 1078 TCAATGGTGAATATTATTACTTCAAGTGGCTTTTAAGCCACAGCAACTATCGGAAG 1137
Qy 363 LysGlnAenThrValSerArgGluArgGlnAenValGluLeuLeuAlaArgGlyArgHis 382
Db 1138 AAGCAAAATACGTGTCAAGGAGCATGAGGATGTTGAACTTTTGGCAAGGGGGCGCAT 1197
Qy 383 AspProCysValAlaProArgAlaValProValValGluSerMetAlaAlaLeuValLeu 402
Db 1198 GACCCCTGTGTGTCCTCGAGCTGTTCTATGTTGGAATCCATGCGCTGCGTGGTCTG 1257
Qy 403 MetAenGlnLeuMetAlaHisValAlaGlnCysGluMetPheAlaLeuAenThrAlaLeu 422
Db 1258 ATGGACCAAGCTCATGGCGCATATTGCCAGTGTGAGATGTTTCGGCTGAACCTTGGCCCTA 1317
Qy 423 GlnGluProValGlySer 428
Db 1318 CAAGAGCCCATTTGGCTCT 1335

```

RESULT 8
US-10-437-963-99207/c
; Sequence 99207, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCES: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 99207
; LENGTH: 1353
; TYPE: DNA
; ORGANISM: Oryza sativa

```

! FEATURE:
! OTHER INFORMATION: Clone ID: PAT_MRT4530_9703C.1
US-10-437-963-99207

Alignment Scores:

```

Pred. No.:      2,71e-185      Length:      1353
Score:          1724.00      Matches:      337
Percent Similarity: 88.04%      Conservative: 31
Best Local Similarity: 80.62%      Mismatches: 32
Query Match:      78.36%      Indels:      18
DB:              19          Gaps:        3

US-10-677-179-8 (1-429) x US-10-437-963-99207 (1-1353)

Qy      15 AlaArgLeuAlaProArgAlaIleGly-----AlaLeuLeuGluPheAlaPro 30
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1251 GCGCGCGTCTCCCGCGCGCGCGCGCGCGCGGTTCGCGGATCCGCGCCG 1192

Qy      31 AlaSerSerLeuArgPheAlaValHisArgCysArgThrAlaArgLeuGluVallys 50
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1191 GCT-----TCCTCCGCTTCCTCGCGCGCGCGCGCGCTCGGCTAGAGGTGAAG 1138

Qy      51 AlaSerGlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGly 70
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1137 GCGTCTGCAAAATGTAATTTGGGAACACTTTCAGGTTGCAACTTATGGAGAGTCTCATGGA 1078

Qy      71 GlyGlyValGlyCysValIleSerGlyCysProProArgIleProLeuThrGluAlaAsp 90
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1077 GCGCGTGTGGTGTGTAATCAGTGGATGCCACCACAGAAATCCCACTTACTGMAAGCAGAT 1018

Qy      91 LeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArglyls 110
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1017 ATGCAGTAGAAGTCACTCGCGGAGCGCGCGCGCGAGAGAGATTAACCCCAAGAAAG 958

Qy      111 GluThrAspThrCysLeuIleLeuSerGlyThrHisGluGlyValThrThrGlyThrPro 130
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      957 GAGACTGACACTTGCAAAATCTTTTCAGGACACATGAAGAAATGACCACTGGGACACCA 898

Qy      131 IleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsn 150
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      897 ATTCAATGTTTGTCCCGCAACACATCAGAGAGGGGGTGATACAGATGNAATGGCTAAG 838

Qy      151 ValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaVal 170
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      837 GCCTACAGACCTTCACATGCAGATCCAACTATGACTTCAATACGGTGTAGAGCAGTG 778

Qy      171 GlnGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeu 190
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      777 CAGCGAGGTGGAAGATCATCAGCAAGAGAGACCATTTGGAAGGGTGGCTGCAGGAGCTCTT 718

Qy      191 ProLysLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysVal 210
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      717 GCMAAGAAATTTCTTAAGCTCAAAATCTGGAGTAGAGATCTTGGCGTGTGTGTCCAAAGTG 658

Qy      211 HisGlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIle 230
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      657 CATCAAGTGTACTACCAAGAGATGCCGTTGATATGACACTGTAAACAAATGGAACAGATA 598

Qy      231 GluSerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIle 250
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      597 GAAGCAACATTTGTTAGATGTCTGTATCCAGAAATATGCACAGAAATGATTGATGCAATC 538

Qy      251 AspArgValArgValArgGlyAspSerValGlyValIleThrCysValAlaArgAsn 270
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      537 GATAAAGTACGAGTTAGAGGTGATTCGATTTGGTGGTGGTGCATCATGATGCAAGAAAT 478

Qy      271 ValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAla 290
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      477 GTTCTCTGGGAGTTGGCTCTCTCTGTATTTGCAAACTTGAGGCTGAATTTGGCGAAAGCT 418

Qy      291 MetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAsp 310
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      417 ATGCTTTCTCTTCTGCAAGCAAGGGGTTTGAGATCGGCGAGTGGATTTGCAGGTACTGAC 358
```

```

Qy      311 LeuThrGlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThr 330
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      357 TACACTGGAAAGTAGCATAATGATGAGTTCTATATGATGAGGTGGAAT----- 307

Qy      331 ArgThrAsnArgSerGlyGlyValGlnGlyIleSerAsnValGluIleValHisPhe 350
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      306 -----GGAGGGATATCAAAATGGTGAATATATATACTTC 274

Qy      351 LysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGlu 370
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      273 AAAGTAGCTTTCAAGCCACCGGACTATTATGGGAAGAAACAACATACTGTTTCAAGGGAG 214

Qy      371 ArgGlnAsnValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaLalaProArgAla 390
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      213 CATGAGGATGTTGAACATTTTAGCAAGGGCGCCACACCAATGTGTTCCTCGCGCT 154

Qy      391 ValProValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisVal 410
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      153 GTGCCGATGTGGAGTCCATGGCGCATTAAGTCTCATGGACGCTGATGGCGCACATT 94

Qy      411 AlaGlnCysGluMetPheAlaLeuAsnThrAlaLeuGlnGluProValGlySer 428
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      93 GCTCAATGTGAGATGTTTCCACTGAACCTTGCTCTACAGAACCAAGTGGCTCT 40

RESULT 9
US-10-425-114-8982
; Sequence 8982, Application US/10425114
; Publication No. US20040034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 8982
; LENGTH: 1674
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700829731_FLI
US-10-425-114-8982

Alignment Scores:
Pred. No.:      4,45e-161      Length:      1674
Score:          1512.00      Matches:      283
Percent Similarity: 82.93%      Conservative: 62
Best Local Similarity: 68.03%      Mismatches: 63
Query Match:      68.73%      Indels:      8
DB:              18          Gaps:        1

US-10-677-179-8 (1-429) x US-10-425-114-8982 (1-1674)

Qy      19 ProArgAlaIleGlyAlaLeuLeuGluPheAlaProAlaSerSer----- 34
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      35 CCATTCTCAGCCGACGCTCTCTCGCGCTTCGCTCTCAATCCGATCCGATCCCTC 94

Qy      35 -----LeuArgPheAlaValHisArgCysArgThrAlaArgLeuGluVallys 50
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      95 CCCCCCGCTACCTCCGACTCTCACTCGCTCTCTCCCAAGACACTTCATACAG 154

Qy      51 AlaSerGlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGly 70
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      155 GCGCTGGAGTACCTACCTGGAATACATTTTCGTTTACACATATGGGAATCATCATGGA 214

Qy      71 GlyGlyValGlyCysValIleSerGlyCysProProArgIleProLeuThrGluAlaAsp 90
```

```
Db 215 GGAGGTGTGGTGTGTTATGATGATGCTCTCGCTCTCTCTCTGAAGCTGAT 274
Qy 91 LeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLys 110
Db 275 ATGCAAGTGGCTCTTGACAGAGAGAGCCAGGTCAGAGCGGAATTACAACCTCTCGAAG 334
Qy 111 GluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThrPro 130
Db 335 GAGACTGATACATGTAATAATATTTTCAGGAGTTTCCGAGGATCTACTACTGGAACCTCA 394
Qy 131 IleLeuValIleValProMetThrAspGlnIleGlySerAspHisArgGluLeuAlaAsn 150
Db 395 ATTCAATGATCTGATCCCAATATCTGATCAAGAGGACATGACTATATGCGAGATGCGAGTA 454
Qy 151 ValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaVal 170
Db 455 GCTTATAGCCCTTCCCATGCAATGCTACTATGATGATGATGATGATGATGATGATGATGAT 514
Qy 171 GlnGlyGlyGlyArgSerGlyArgLysThrValGlyValAlaAlaAlaGlyAlaLeu 190
Db 515 CAGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 574
Qy 191 ProLysLysIleLeuLysLysLysCysGlyLeuGluLeuLeuSerPheValSerLysVal 210
Db 575 GCTAAGAAATCTTAAAGAAATTTCTCGAACTGAGATCTCGCCTATGCTCTCAAGTT 634
Qy 211 HisGlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIle 230
Db 635 CATAAGATTTGTTTCCAGAGACCTTATGATCATGACACTCTGACTCTTCATCAGATT 694
Qy 231 GluSerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIle 250
Db 695 GAGAGTAACATTTGTCGATGTCAGACCCGAGTATGACAGAGAATGATATCTGCAATT 754
Qy 251 AspArgValArgValArgGlyAspSerValGlyGlyValIleThrCysValAlaAlaArgAsn 270
Db 755 GATGCTGTGCGAGTGAGAGTGATCTGTTGTGGTGTGTTGTGACATGCATGTCATGAGGAAC 814
Qy 271 ValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAla 290
Db 815 TGTCACGAGGTCGCGTTTCCAGCATTTTGAACAACTTGAAGCTGAGCTGGCTAAAGCT 874
Qy 291 MetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAsp 310
Db 875 GCATGTCATTCCTGTCACCAAGGCTTTCAAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 934
Qy 311 LeuThrGlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThr 330
Db 935 TTGACTGGAGTGAACACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 994
Qy 331 ArgThrAsnArgSerGlyValGlnGlyGlyIleSerAsnValGluIleValHisPhe 350
Db 995 AGAACAAATTCGCTCTGGTGGGATCAGGGTGGAAATTCCTCAATGGGGAATCAATTAATATG 1054
Qy 351 LysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGlu 370
Db 1055 AGAATAGCTTTCACGCCAACATCAACATTTGGAAGAGCAAAAGACTGTGACTCGAGAT 1114
Qy 371 ArgGlnAsnValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaProArgAla 390
Db 1115 AAAAAAGAACACAGAGTTTATAGCCCTGGTGGCCATGATCTTGTGTTGTTGCCAAGAGCT 1174
Qy 391 ValProValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisVal 410
Db 1175 GTACCTATGGTAGAAGCAATGTTAGCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1234
Qy 411 AlaGlnCysGluMetPheAlaLeuAsnThrAlaLeuGlnGluProVal 426
Db 1235 GCGCAGTGTATCTTTTCCCGTAACACTCAGATTTTGCAGAACCCCTTG 1282
```

RESULT 10

US-10-424-599-1811

```
/ Sequence 1811, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa Thomas J
/ APPLICANT: Kovalic David K
/ APPLICANT: Zhou Yihua
/ APPLICANT: Cao Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 1811
/ LENGTH: 2094
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_101634C.1
US-10-424-599-1811
```

```
Alignment Scores:
Pred. No.: 6,1e-161 Length: 2094
Score: 1512.00 Matches: 283
Percent Similarity: 82.93% Conservative: 62
Best Local Similarity: 68.03% Mismatches: 63
Query Match: 68.73% Indels: 8
DB: 18 Gaps: 1
```

US-10-677-179-8 (1-429) x US-10-424-599-1811 (1-2094)

```
Qy 19 ProArgAlaIleGlyAlaLeuLeuGluPheAlaProAlaSerSerSer----- 34
Db 220 CATTCTCAGCGGAGCTCTCTCGCCCTTCGCTCTCTCAATTCGATTCGGATCCCTC 279
Qy 35 -----LeuArgPheAlaValHisArgCysArgThrAlaArgLeuGluValLys 50
Db 280 TCCCCCGCTACCTCCGACTCTCACTCGCTCGCTCTCCCAAGAGACTTCGCATACAG 339
Qy 51 AlaSerGlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGly 70
Db 340 CGCGCTGGGAGTACCTGGAATCACTTCGTGTACAAATATGCGGGAATCACAATGGA 399
Qy 71 GlyGlyValGlyCysValIleSerGlyCysProProArgIleProLeuThrGluAlaAsp 90
Db 400 GGAGGTGTGGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 459
Qy 91 LeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLys 110
Db 460 ATGCAAGTGGATCTTGACAGAAGGAGGCGCAGTCCAGAGCCGAATTACAACCTCTAGAAG 519
Qy 111 GluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThrPro 130
Db 520 GAGACTGATACATGTAATAATATTTTCAGGAGTTCCTCGAAGGAATCACTACTCTGGAAC 579
Qy 131 IleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluLeuAlaAsn 150
Db 580 ATTCAATGATCTGATCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 639
Qy 151 ValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaVal 170
Db 640 GCTTATAGCCCTTCCCATGCAATGCTACTATGATGATGATGATGATGATGATGATGATGATGAT 699
Qy 171 GlnGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeu 190
Db 700 CAGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 759
Qy 191 ProLysLysIleLeuLysLysLysCysGlyLeuGluLeuLeuSerPheValSerLysVal 210
Db 760 GCTAAGAAATCTTAAAGAAATTTTCTGGAACTCAGATTTCTGGCCTATGCTCTCAAGTT 819
Qy 211 HisGlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIle 230
```

US-10-696-616-27

Alignment Scores:

Pred. No.:	1-23e-154	Length:	1311
Score:	1454.00	Matches:	268
Percent Similarity:	84.09%	Conservative:	65
Best Local Similarity:	67.68%	Mismatches:	63
Query Match:	66.09%	Indels:	0
DB:	19	Gaps:	0

US-10-677-179-8 (1-429) x US-10-696-616-27 (1-1311)

Qy	31	AlaSerSerSerLeuArgPheAlaValHisArgCysArgThrAlaArgLeuGluValLys	50
Db	88	TCVTTCTCCGCCGCTTCAGATCTCTCTCCGTACCAAAACAGGAAGAACTTCCAGATACAA	147
Qy	51	AlaSerGlyAsnThrPheGlyAsnThrPheGlnValAlaThrTyGlyGluSerHisGly	70
Db	148	GCTACTCGGAAGTTCAVATGGGACTCAATTTTCGAGTTTCAACTTTTGGAGAATCACATGGGA	207
Qy	71	GlyGlyValGlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAsp	90
Db	208	GGAGGAGTTGGTTGTATCATGTATGGTTGCTCTCTCGTATTCACATTACTGAACTTCAT	267
Qy	91	LeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLys	110
Db	268	TTGCAATTGCATCTCGATAGAGGAGCGCTGGTCAGACGAGGATCACAACTCCTTAGAAAA	327
Qy	111	GluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThrPro	130
Db	328	GAGACTGATACTTGGCGGATATCGTCTGGAGTCTCTGAAGGAATGACGACAGGAACACCT	387
Qy	131	IleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsn	150
Db	388	ATCATGTGTTGTATCCAAACACAGATCAGAGAGGACTTGATTTACAGTGAATGTGCGGTT	447
Qy	151	ValTyArgProSerHisAlaAspAlaThrTyAspPheLysTyGlyValArgAlaVal	170
Db	448	GCCTATAGACCATCGCATGCTGTATGCAACTTATGACATGAAGTATGTTGTTCAGATCAGT	507
Qy	171	GlnGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeu	190
Db	508	CAGGTTGAGGAGAGATCTTCAGCTAGAGACACATTGGAAAGAGTTGCTCTCGGAGCTTTG	567
Qy	191	ProLysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysVal	210
Db	568	GCCAAAGAAATTTGAAGCAATTTGCAGGAACCTGAGATTTCTGCTATGTTCTCGCAAGTT	627
Qy	211	HisGlnValValLeuProGluAspAlaValAspTyArgLysSerValThrLeuGluGlnIle	230
Db	628	CACCATGTTGTACTTCCAGAAGAATTTGGTAGACCACGAGAATTTAACACTCGAACAGATA	687
Qy	231	GluSerAsnIleValArgCysProAspProGluTyAlaGluLysMetIleAspAlaIle	250
Db	688	GAATAATACATTTGTTCAGATCCCTTAATCCCGAGTATCGGAAAGAGATGATAGCTGCGATT	747
Qy	251	AspArgValArgValArgGlyAspSerValGlyValIleThrCysValAlaArgAsn	270
Db	748	GATGCTGTCCAGGCAAAAGGGAACCTGTTGGTGGTGTGTGACCTGCATTTGTCGGAAT	807
Qy	271	ValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAla	290
Db	808	GCTCCACGTGGGCTTGGTACACCGGTTTTTCGATAAACTTGAAGCAGAACTGGCAAAAGCT	867
Qy	291	MetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAsp	310
Db	868	TGATGTGCGTACCTGCAACTTAAGGGGATTTGAGTTTGGAAACGCGCTTTGCGAGTACCTTT	927
Qy	311	LeuThrGlySerGluHisAsnAspGluPheTyMetAspLysAlaGlySerValArgThr	330
Db	928	TTGACTGGTCTTGAAACAATGATGAGTCTTATACCGATGAATGAAGNATACGTACCT	987
Qy	331	ArgThrAsnArgSerGlyValGlnGlyLysSerAsnValGluIleValHisPhe	350

```
Db 988 AGAACCAACCGATCTGGTGAATTCAGGAGGATCTCAATGTGTAATAATAAACATG 1047
Qy 351 LysValAlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGlu 370
Db 1048 AGAGTAGCCTTCAAGCAACATCAACAAATGGGAAGCAAAATACGGTAACCGAGAG 1107
Qy 371 ArgGlnAsnValGluLeuAlaArgGlyArgHisAspProCysValAlaProArgAla 390
Db 1108 AAGGTAGAACCGAATGATGGCGGTGGTGCATGATCTTGTTGTTCCAGAGCT 1167
Qy 391 ValProValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisVal 410
Db 1168 GTGCCAATGGTGAAGCAATGGTGGCTCTAGTTCTTGTGGATCAATGTATGGCGCAATAC 1227
Qy 411 AlaGlnCysGluMetPheAlaLeuAsnThrAlaLeuGlnGluProVal 426
Db 1228 GCACCAATGCCATTTGTTCCAAATAATCCAGAGTTTCAGGAACCTCTC 1275

RESULT 12
US-10-677-179-11
; Sequence 11, Application US/10677179
; Publication No. US20040082050A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Pember, Stephen O.
; TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
; FILE REFERENCE: BB-1159-A
; CURRENT APPLICATION NUMBER: US/10/677,179
; PRIOR FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: US/09/743,207
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: PCT/US99/16353
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/093,611
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-677-179-11

Alignment Scores:
Pred. No.: 1,93e-132 Length: 966
Score: 1257.00 Matches: 245
Percent Similarity: 89.30% Conservative: 22
Best Local Similarity: 81.94% Mismatches: 26
Query Match: 57.14% Indels: 6
DB: 18 Gaps: 2

US-10-677-179-8 (1-429) x US-10-677-179-11 (1-966)
Qy 15 AlaArgLeuAlaProArgAlaIleGly-----AlaLeuGluPheAlaPro 30
Db 75 GCGCGGCTCTCCCGCGCGCGCGCGCGGCGGTTCGCGCCTTCCGAGTCCGCCCG 134
Qy 31 AlaSerSerLeuArgPheAlaValHisArgCysArgThrAlaArgLeuGluVallys 50
Db 135 GCT-----TCCTCCGCTCTCCGTCGCGCGCGCGCGCGCTGCTAGAGGTGAAG 188
Qy 51 AlaSerGlyAsnThrPheGlyAsnThrPheGlnValAlaThrTyrGlyGluSerHisGly 70
Db 189 GCGTCTGCAAAATGTAATTTGGGAACCTACTTCCAGGTTGCAACTTATGGAGAGTCTCATGGA 248
Qy 71 GlyGlyValGlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAsp 90
Db 249 GCGCGGTGTGGTGTGATAGTGATGCGCCAGCAATCCCACTTACTGAAGCAGAT 308
Qy 91 LeuGlnValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArglys 110
Db 110 GCGCGGTGTGGTGTGATAGTGATGCGCCAGCAATCCCACTTACTGAAGCAGAT 308
```

```
Db 309 ATGCAAGTAGAACTCGACCGGAGCGCCAGCAGAGCAAGTAACACCCCAAGAAAG 368
Qy 111 GluThrAspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrClyThrPro 130
Db 369 GAGACTGACACTTGCANAATCTTTTCAGGACACATGAAGGAATGACCATCGGACACCA 428
Qy 131 IleLeuValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsn 150
Db 429 ATTCATGTTTGTTCCTCCGACACATGAGAGGGGGGTGATTACAGTGAAATGCGCTAAG 488
Qy 151 ValTyrArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaVal 170
Db 489 GCGTACAGACCTTTCATGTCAGATGCAACTTATGACTTCAATAACGTTGTAGACAGTG 548
Qy 171 GlnGlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeu 190
Db 549 CAGGAGGTGGAAGATCATCAGCAAGAGAGACCATTTGGAAGGGTGGCTGCAGGAGCTCTT 608
Qy 191 ProlLysIleLeuLysLysCysGlyLeuGluIleLeuSerPheValSerLysVal 210
Db 609 GCMAAGAAATTCCTTACGCTCAATCTGGAGTAGAGATCTTGGCGTTTGTGTCGAAGGTG 668
Qy 211 HisGlnValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIle 230
Db 669 CATCAAGTTGTACTACCAAGAGATGCCGTTGATTATGACACTGTAAACAATGGAACAGATA 728
Qy 231 GluSerAsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIle 250
Db 729 GAAAGCAACATTTGTAGATGCTCCTGATCCAGAAATATGCACAGAAGATGATTGATGACTC 788
Qy 251 AspArgValArgValArgGlyAspSerValGlyGlyValIleThrCysValAlaArgAsn 270
Db 789 GATAAAGTACAGTTAGAGGTGATTCGATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 848
Qy 271 ValProArgGlyLeuGlySerProValPheAspLysLeuGluSerGlyLeuAlaLysAla 290
Db 849 GTTCTCGTGGATGGCTCTCTCTGATTTTGCACAACTTGAGGCTGAATTCGCGAAGCT 908
Qy 291 MetLeuSerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThr 309
Db 909 ATGCTTTCTCTCTGCAAGCAAGGGGTTTGAGATCGCAGTGGATTGTGTGTTCACT 965

RESULT 13
US-10-369-493-26364
; Sequence 26364, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 26364
; LENGTH: 1089
; TYPE: DNA
; ORGANISM: Synecocystis sp.
US-10-369-493-26364

Alignment Scores:
Pred. No.: 6,32e-125 Length: 1089
Score: 1191.50 Matches: 225
Percent Similarity: 77.90% Conservative: 57
Best Local Similarity: 62.15% Mismatches: 77
Query Match: 54.16% Indels: 3
DB: 17 Gaps: 1
```



```
US-10-677-179-8 (1-429) x US-10-369-493-42555 (1-1089)
Qy 53 GlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGlyGlyGly 72
Db 4 GAAACACGTTTGGCAGTTTATTTTCGCATTAACCACTTCGCGTGGAGTCCATGGGGGGGC 63
Qy 73 ValGlyCysValIleSerGlyCysProProArgIleProLeuThrGluAlaAspLeuGln 92
Db 64 GTGGGGGTAATATTATGATGGTTGGCTCCCGATTTGGAAATTTCCCCAGAGGAATCCAA 123
Qy 93 ValGluLeuAspArgArgProGlnSerArgIleThrSerThrArgIleGlyLeuThr 112
Db 124 GTAGACCTAGACCGCGCGGCGGACGACAGCAAAATTAACACCCCTCGTAAGGAAGCA 183
Qy 113 AspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThrProIleLeu 132
Db 184 GATCAATCGAAATTTCTCTGGCGTGTTCGAGGGTAAACTCTGGGTACACCGATCGCC 243
Qy 133 ValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnValTyr 152
Db 244 ATTTTGGTGGAAATAAGATGCCCTTCACGATTAACAATGAGATGCGGTGAAATAT 303
Qy 153 ArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGlnGly 172
Db 304 CGTCTCTCCACGCGCATGCCACCTATGACGGAATATGCGATTCGCAATGGCAAGGA 363
Qy 173 GlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuProLys 192
Db 364 GGAGGTGCGCTCTTCTGTAGGAAACCACTTGGCGGGTGGCGGAGGGCGGATGCCAAA 423
Qy 193 LysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysValHisGln 212
Db 424 AAGATTCTGGCCCAAGTTTAACGGGTGGAAATTTGGGCTATGTGAAAGTATCAAGAC 483
Qy 213 ValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGluSer 232
Db 484 ATT-----GAGCGCGGTGGACAGCAATACGCTGACCTTGGACAGTGGAAAGC 534
Qy 233 AsnIleValArgCysProAspProGluTyrAlaGluLysMetIleAspAlaIleAspArg 252
Db 535 AATAATTGTTGTTGTCGTCGCGATGAAGATGTGCCGAAAAAATGATTTGAACGCGATGAC 594
Qy 253 ValArgValArgGlyAspSerValGlyValIleThrCysValAlaAlaArgAsnValPro 272
Db 595 GTGTACGACAAAAAGATCTCATTTGGCGGGTGGTGGAAATGCGCATCCGTAACGCTCCC 654
Qy 273 ArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMetLeu 292
Db 655 AAGGGTGGGGGAACCAAGTGTGTTGATTAAGTTGGAGCGGATTTGGCCAGGCAATGATG 714
Qy 293 SerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeuThr 312
Db 715 TCCCTCCCGCTTACCAGGGGTTTCAGATTTGGCTCTGGTGTTCGGGTACTTTACTGACT 774
Qy 313 GlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThrArgThr 332
Db 775 GGTAGTCAACATAATGATGAGTACTTACCTGGATGAAGCGGGGAATGGGTACTTCGCACC 834
Qy 333 AsnArgSerGlyGlyValGlnGlyIleSerAsnValGluIleValHisPheLysVal 352
Db 835 AACCGTCTGGGGGGTTTCAGGAGGCGATCAGCAACGCGGAACCCATTTATTCGCATA 894
Qy 353 AlaPheLysProThrProSerIleGlyValLysGlnIleAsnThrValSerArgGluArgGln 372
Db 895 GCATTTAAACCTACTTCGCCACCACTTGGTCAAGAGCAAAAAACCGTTAGTAACATGGGGG 954
Qy 373 AsnValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaProArgAlaValPro 392
Db 955 GAAACCACTTAGCCGCTAAAGTGCCTATGATCCCTCGTTTCGCCCGGCTGTGCC 1014
Qy 393 ValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisValAlaGln 412
```

```
Db 1015 ATGCTGGAAGCGATGGCTGCATTAGTACTTTGTGACCACTTTGTTCGATTCCAAGCCAG 1074
Qy 413 CysGlu 414
Db 1075 TGTAAA 1080
RESULT 14
US-10-369-493-42555
; Sequence 42555, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 42555
; LENGTH: 1086
; TYPE: DNA
; ORGANISM: Anabaena PCC7120
US-10-369-493-42555
Alignment Scores:
Pred. No.: 2,33e-124 Length: 1086
Score: 1186.50 Matches: 220
Percent Similarity: 78.79% Conservative: 66
Best Local Similarity: 60.61% Mismatches: 74
Query Match: 53.93% Indels: 3
DB: 17 Gaps: 1
US-10-677-179-8 (1-429) x US-10-369-493-42555 (1-1086)
Qy 53 GlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGluSerHisGlyGlyGly 72
Db 4 GCAATACTTTTGGACATCTGTTTCGGATCCTACTTTTGGCGAGTCACACGGTGGAGGC 63
Qy 73 ValGlyCysValIleSerGlyCysProProArgIleProLeuThrGluAlaAspLeuGln 92
Db 64 GTGGGGTGTGTGATGTTGACGCTGTCTCCGCTACTCGAAATTTCTCTCGAGGAAATTCAA 123
Qy 93 ValGluLeuAspArgArgProGlyGlnSerArgIleThrSerThrArgLysGluThr 112
Db 124 CTGAGATTAGCCGACAGACCCGACAGAGTAAGATTACAGCCCTCGTAAGGAAGCA 183
Qy 113 AspThrCysLysIleLeuSerGlyThrHisGluGlyValThrThrGlyThrProIleLeu 132
Db 184 GACACCTGTGAGATTTTGTCTGGGGTATATGAAGGCAAAACTCTGGGAACGCCCAATATCA 243
Qy 133 ValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnValTyr 152
Db 244 ATTTTGTGCGAAATAAGATATCTGTCGCCAGATTTATGATGAGATGGCGCAAAAGTAT 303
Qy 153 ArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGlnGly 172
Db 304 CGTCTTCCCATGCTGATGCAACTTATGATGCTTAATATGGGATTCGTAATTCGCAAGGT 363
Qy 173 GlyGlyArgSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuProLys 192
Db 364 GGCGGTAGATCTCTCAGCGCGGAGACAATTTGGTAGAGTAGCCGCGAGTGGCGATCGCTAAA 423
Qy 193 LysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysValHisGln 212
Db 424 AAAATCTCCGCAAGTTGCTAATGTAGAGTATATGAGTACGCTCAACCGTATCAAGAC 483
Qy 213 ValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGluSer 232
```



```

Db      484  TTG-----GAAGGTGTTGTAGATCCCAATACCGTCACCTAGACCAAGTGGAAAGC 534
Qy      233  AenlleValAtqCysProAspProGluTyrAlaGluLysMetIleAspAlaIleAspArg 252
Db      535  AATATCGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 594
Qy      253  ValArgValArgGlyAspSerValGlyGlyValIleThrCysValAlaAlaArgAsnValPro 272
Db      595  ACAGGTAGACAAAGGTGATCTATCGCGGTGTAGTGGATGGATGGATGGATGGATGGAT 654
Qy      273  ArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMetLeu 292
Db      655  AAAGCGTGGGTGAACCTGATTCATCGCGGTGTAGTGGATGGATGGATGGATGGATGGAT 714
Qy      293  SerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeuThr 312
Db      715  TCTTACCTGCTAGCAAGGTTTGAATCGGTTTCAAGGTTTCAAGGTTTCAAGGTTTCAAG 774
Qy      313  GlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThrArgThr 332
Db      775  GGATTTGACATAACGACGATATATATATATATATATATATATATATATATATATATAT 834
Qy      333  AsnArgSerGlyValGlnGlyGlyIleSerAsnValGluIleValHisPheLysVal 352
Db      835  AACCGTTCGGTGGGATTCAAGCGGTATTGCAACGCGAGAGAAATATCATTTTGGAGTT 894
Qy      353  AlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGluArgGln 372
Db      895  GCATTTAAACCCACACGACGATAGAAAGAGCAAAACAGTTACTCTGTAAGCGCAA 954
Qy      373  AsnValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaAlaProArgAlaValPro 392
Db      955  GAAACCTATTAGCGCCCAAGGTAGACATGATCTCTGTTCTACCCCGCGAGTCCCG 1014
Qy      393  ValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisValAlaGln 412
Db      1015  ATGTTAGAAAGCAATGGTGGCGTGTGCTACTGTGCGATCATTTGTTAGCGCATCATGGACAG 1074
Qy      413  CysGluMet 415
Db      1075  TGTAAAGGTG 1083

```

RESULT 15

```

US-10-369-493-43897
; Sequence 43897, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 43897
; LENGTH: 1086
; TYPE: DNA
; ORGANISM: No. US20030233675A1toc punctiforme
US-10-369-493-43897

```

Alignment Scores:

Pred. No.:	6.63e-124	Length:	1086
Score:	1182.50	Matches:	220
Percent Similarity:	78.51%	Conservative:	65
Best Local Similarity:	60.61%	Mismatches:	75
Query Match:	53.75%	Indels:	3

```

DB:      17      Gaps:      1
US-10-677-179-8 (1-429) x US-10-369-493-43897 (1-1086)
Qy      53  GlyAsnThrPheGlyAsnTyrPheGlnValAlaThrTyrGlyGlySerHisGlyGlyGly 72
Db      4  GGCATAATTTTGGTTCATTTATTTTCGATTAGTACTTTTGGCGAGTCTTCACGGCGCGGT 63
Qy      73  ValGlyCysValIleSerGlyCysProArgIleProLeuThrGluAlaAspLeuGln 92
Db      64  GTGGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 123
Qy      93  ValGluLeuAspArgArgArgProGlyGlnSerArgIleThrSerThrArgLysGluThr 112
Db      124  GTAGAACTAGATAGAGCGCGCGGCAAAAGTAAATACGACTCCCGCCCAAGAAGCT 183
Qy      113  AspThrCysLysIleLeuSerGlyThrHisGluGlyValThrGlyThrProIleLeu 132
Db      184  GATACCTCGAGATTATCTCAGGAGTATTTGAAGCAAAACGCTGGGAACCCCTATAACG 243
Qy      133  ValIleValProAsnThrAspGlnIleGlySerAspHisArgGluIleAlaAsnValTyr 152
Db      244  ATTTTGTGTACGTAATCAAGATACCTCGTCCCAAGATTATGACGAGATGGCACAGAAGTAT 303
Qy      153  ArgProSerHisAlaAspAlaThrTyrAspPheLysTyrGlyValArgAlaValGlnGly 172
Db      304  CGGCTTCTCAGCGGATGCACCTATGATGCAAAATATGCGCATTCGCAATTGGCAAGGT 363
Qy      173  GlyGlyArgSerSerGlyArgLysThrValGlyArgValAlaAlaGlyAlaLeuProLys 192
Db      364  GGGGTAGGTCTGTCAGCACGTCAGACCAATCGGAGAGTAGCAGCAGGTGCGATCGCTAAA 423
Qy      193  LysIleLeuLysLeuLysCysGlyLeuGluIleLeuSerPheValSerLysValHisGln 212
Db      424  AAAATTCCTCGCTCAAGTTGCCAATGTCGAAATTTATGCTTACGTTAAGCGGATTAAGAC 483
Qy      213  ValValLeuProGluAspAlaValAspTyrGlySerValThrLeuGluGlnIleGluSer 232
Db      484  TTG-----GAAGGTGTAGTGCATCTTAATATCTGTAACCTTAGAACACAGTAGAAGC 534
Qy      233  AsnIleValArgCysProAspProGluTyrAlaGlyLysMetIleAspAlaIleAspArg 252
Db      535  AATATCGTTCGCTGCTCGCATGCTGAATGTGGCGATGCGCATGATGAATGATTGATTGAG 594
Qy      253  ValArgValArgGlyAspSerValGlyValIleThrCysValAlaAlaArgAsnValPro 272
Db      595  ATAGGTAGACAAAGCGGATTCATCGCGGTGTGCTAGAAATGTGTGGCGCGAAATGTACCG 654
Qy      273  ArgGlyLeuGlySerProValPheAspLysLeuGluSerGluLeuAlaLysAlaMetLeu 292
Db      655  AAAGTTTGGCGCAACCAAGTATTTGATATAATTAGAACTGATATCGCTTAGGGTGTCTAG 714
Qy      293  SerIleProAlaSerAsnGlyPheGluIleGlySerGlyPheAlaGlyThrAspLeuThr 312
Db      715  TCTCTGCTGCTAGCAAAAGGCTTTGAAATTTGTTTCCGTTTTCGGGAAACGCTACTAACG 774
Qy      313  GlySerGluHisAsnAspGluPheTyrMetAspLysAlaGlySerValArgThrArgThr 332
Db      775  GGAATTGAGCATACGACGAAATTTATATATATATATATATATATATATATATATATATAT 834
Qy      333  AsnArgSerGlyGlyValGlnGlyIleSerAsnValGluIleValHisPheLysVal 352
Db      835  AACCGTTCGGCGTATTCAAGCGCGGATTTCTTACCGCGCAAAATATCATTTTTCGAGTT 894
Qy      353  AlaPheLysProThrProSerIleGlyValLysGlnAsnThrValSerArgGluArgGln 372
Db      895  GCATTTAAGCGCACGACCAATTTAGAAAAGACAGAAAGACTGTAACTCGTGGGCGCAA 954
Qy      373  AsnValGluLeuLeuAlaArgGlyArgHisAspProCysValAlaAlaProArgAlaValPro 392
Db      955  GAAACCTATTAGCAGCAAAAGGACGACGATCTTTGTGTATTACCGCGTGCAGTGCACA 1014
Qy      393  ValValGluSerMetAlaAlaLeuValLeuMetAspGlnLeuMetAlaHisValAlaGln 412

```

Db 1015 ATGTTGAGGCAATGGTGGCGTTGGTACTATGTGACCAATTGTTACGGCATCATGGTCAG 1074
QY 413 CysGluMet 415
Db 1075 TCGAAAGTC 1083

Search-completed: August 27, 2005, 07:26:24
Job time : 755 secs

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☐ BLACK BORDERS

☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES

☐ FADED TEXT OR DRAWING

☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING

☐ SKEWED/SLANTED IMAGES

☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS

☐ GRAY SCALE DOCUMENTS

☐ LINES OR MARKS ON ORIGINAL DOCUMENT

☒ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)